Skill Management in Large-scale Service Marketplaces

Gad Allon
Kellogg School of Management, Northwestern University, Evanston, Illinois, g-allon@kellogg.northwestern.edu

Achal Bassamboo
Kellogg School of Management, Northwestern University, Evanston, Illinois, a-bassamboo@kellogg.northwestern.edu

Eren B. Çil
Lundquist College of Business, University of Oregon, Eugene, Oregon, erencil@uoregon.edu

Large-scale, web-based service marketplaces have recently emerged as a new resource for customers who need quick resolutions for their short-term problems. Due to the temporary nature of the relations between customers and service providers (agents) in these marketplaces, customers may not have an opportunity to assess the ability of an agent before their service completion. On the other hand, the moderating firm has a more sustained relationship with agents, and thus it can provide customers with more information about the abilities of agents through skill screening mechanisms. In this paper, we consider a marketplace where the moderating firm can run two skills tests on agents to assess if their skills are above certain thresholds. Our main objective is to evaluate the effectiveness of skill screening as a revenue maximization tool. We, specifically, analyze how much benefit the firm obtains after each additional skill test. We find that skill screening leads to negligible revenue improvements in marketplaces where agent skills are highly compatible. As the compatibility of agent skills weakens, we show that the firm starts to experience as much as 25% improvement in revenue from skill screening. Apparently, the firm can reap the most of these substantial benefits when it runs only one test. For instance, in marketplaces where agents posses uncorrelated skills, the second skill test only brings an additional 1.7% improvement in revenue. Accounting for possible skill screening costs, we then show the optimality of offering only one test when the compatibility between agent skills is sufficiently low. The results of this paper also have important implications in terms of the right level of intervention in the marketplaces we study.

Key words: Service marketplaces; skill management; flexible resources; price competition; non-cooperative game theory.

---

1. Introduction

Large-scale, web-based service marketplaces have recently emerged as a new resource for customers who need quick resolutions for their temporary problems. In these marketplaces, many small service providers (agents) compete among themselves to help customers with diverse needs. Typically, an independent firm, which we shall refer to as the *moderating firm*, establishes the infrastructure for the interaction between customers and agents in these marketplaces. In particular, the moderating firm provides the customers and the agents with the information required to make their decisions. A notable example among many existing online marketplaces is upwork.com (formerly odesk.com). The web-site hosts around 9,000,000 programmers competing to provide software solutions.

Considering large-scale natures of online marketplaces, it is not surprising to see that the ability of agents to serve customers with a particular need varies significantly. Naturally, customers prefer to be serviced by a more skilled agent because a more capable agent is likely to generate more value for customers. Unfortunately, customers may not have an opportunity to assess the ability of an agent before their service completion because most of the relations between customers and providers are temporary in these marketplaces. On the other hand, the moderating firm has a more sustained relationship with agents, and thus it can obtain more information about their abilities. Particularly, the firm can constitute a *skill screening mechanism*. In general, these mechanisms take the form of skill tests and/or certification programs that are run by moderating firms. For instance, upwork.com offers various exams to test the ability of the *candidate* providers. In fact, being successful in some of these exams is the first requirement for providers to be eligible to serve customers in the marketplace. The makers of upwork.com (or any other moderating firm) freely decide on how comprehensive the exams are. The more comprehensive the exams become, the more value customers expect from their service. If necessary, upwork.com can use these exams to disqualify some of the agents, and thus control the portfolio of different agent types (e.g., flexible, dedicated) and the service capacity in the marketplace. We use the term *skill-mix structure* to denote the portfolio of different agent types. Most of the online service marketplaces, including upwork.com, receives 10% of the revenue obtained by the agents at service completion. Therefore, it is in the best interest of the moderating firm to intervene in the marketplace by using its skill tests in order to make sure that the “right” prices and customer demand emerge in the marketplace.

Motivated by these online service marketplaces, we consider a marketplace with two groups of customers, each of which has different needs. We use the term *class* to identify the group of customers with similar needs. On the supply side, we assume there is a population of agents who are homogeneous in their service capacity and heterogeneous in the value that their services generate for each customer class. Specifically, the service of an agent generates a random value (with a known distribution) for each customer class. We refer to the vector that represents these two random values as *agent skills*. Customers cannot observe agent skills, but the moderating firm can run two skills tests on agents to assess whether their skills are above certain thresholds or not. The firm determines a passing level in each skill test and allows agents to serve customers only if their skills are above the passing levels. We use a game theoretical framework to study the interaction between the customers and the agents. Namely, each agent announces a price for his service, and customers request service from agents based on the expected agent skills, the prices, and the likelihood of being served. The objective of the firm is to find the passing levels that maximize its revenue, which is a predetermined share of the total revenue generated in the marketplace.

As we mention above, the moderating firm can use its screening mechanism as a tool to influence the revenue generated in the marketplace. In this paper, we aim at evaluating the effectiveness of skill screening as a revenue maximization tool. To this end, we analyze how much benefit the firm obtains after each additional skill test. We start with a benchmark case where the firm does not offer any skill tests. Then, we study the firm's problem (i) when the firm uses only one exam, and (ii) when both exams are offered. Due to test preparations and executions, moderating firms may incur a cost to carry out the skill tests. Taking the possibility of costly skill screening into consideration, we compare the revenues in these three cases to find the optimal number of skill tests. One may view the number of tests offered by the firm as a measure for how strictly the firm regulates the marketplace. Thus, our analysis also provides insights for the right level of intervention in the marketplaces we study.

In analyzing the model described above, we observe that the optimization problem of the firm becomes analytically intractable when the skills of an agent follow a general joint probability distribution. Thus, we obtain the firm's optimal decisions by considering a family of skill distributions with a shape parameter which controls the correlation between agent skills. Since the online marketplaces we review usually house service providers with compatible skills, we focus on positively correlated skills in this paper.

Our model of agent skills enables us to gain a better understanding of the relationship between skill correlation and how the firm utilizes skill screening. We, specifically, study how the revenue benefits from skill screening depend on the correlation between the agent skills. We find that skill screening leads to minimal revenue improvements in marketplaces where agent skills are highly correlated. If the moderating firm is concerned about the cost of screening, this result suggests that the firm is better off not offering any skill tests when the skill correlation is high. As the compatibility of agent skills weakens, we show that the firm starts to experience substantial revenue benefits from skill screening. Particularly, we prove that the firm's benefit from skill screening can be as much as 25% if agent skills exhibit negligible correlation. Apparently, the firm can reap the most of these substantial benefits when it runs only one exam. For instance, in marketplaces with almost independent skills, the second skill test can only bring an additional 1.7% revenue improvement. Accounting for possible testing costs, we then show the optimality of offering only one test when skill correlation is sufficiently low.

The results of this paper also have important implications in terms of the moderating firm's involvements in the marketplace. Our findings suggest that the firm does not need to regulate the marketplace via skill screening when agents are endowed with highly compatible skills. When intervention is needed, we establish that it is sufficient to run only one of the exams as an intervention tool when considering costly skill screening. The contribution of this paper is also in introducing a family of joint skill distributions that captures the correlation between skills ranging from perfect and positive correlation to no correlation. Our methodology can be easily extended to study the service environments with negatively correlated skills.

2. Literature Review

Our paper lies in the intersection of various streams of research. The first line of works related to our paper studies customer behavior in service systems. Service systems with customers who seek to maximize their utilities have attracted the attention of researchers for many years. The analysis of such systems dates back to Naor's seminal work (See Naor (1969)), which analyzes customer behavior in a single-server queueing system. More recently, Cachon and Harker (2002) and Allon and Federgruen (2007) study the competition between multiple firms offering substitute but differentiated services by modeling the customer behavior implicitly via an exogenously given demand function. An alternative approach is followed in Chen and Wan (2003), where authors examine the customers' choice problem explicitly by embedding it into the firms' pricing problem.

Our paper is also related to the research focusing on the economic trade-offs between investing on flexible resources, which provide the ability to satisfy a wide variety of customer needs, and dedicated resources responding to only a specific demand type. This line of literature studies a two-stage decision problem with recourse, which is also known as the Newsvendor Network problem, and dates back to Fine and Freund (1990). Fine and Freund (1990) considers a firm that invests in a portfolio of multiple dedicated resources and one flexible resource in the first stage where the market demand for its products is uncertain. After making the capacity investments, the demand uncertainty is resolved, and the firm makes the production decisions to maximize its profit. Fine and Freund (1990) argues that the flexible resource is not preferred when demand distributions are perfectly and positively correlated. Gupta et al. (1992) studies a similar model where the firm initially has some existing capacity and presents results parallel to Fine and Freund (1990). Contrary to the examples provided in these two papers, Callen and Sarath (1995) and Van Mieghem

(1998) show that it can be optimal for a firm to invest in a flexible resource even if demand distributions are perfectly and positively correlated. Recent papers extend the model in Fine and Freund (1990) by studying the optimal pricing decision of a monopolist (See Chod and Rudi (2005) and Bish and Wang (2004)), competition between two firms (See Goyal and Netessine (2007)), and more detailed configurations of flexibility (See Bassamboo et al. (2010)). In all of these papers, the firm chooses its price and allocates its flexible capacity in order to maximize its profit. However, in the service marketplaces we consider, the (moderating) firm does not have direct control over the pricing and the service decisions of the service providers.

To our knowledge, Allon and Gurvich (2010) and Chen et al. (2008) are the first papers studying competition among complex service systems with many independent decision makers. There are two main differences between these two papers and our work. First, both of them study a service environment with a fixed number of decision makers (firms), while the number of decision makers in our marketplace (agents) is infinite. Second, they only consider a competitive environment where the firms behave individually. In contrast, we study a marketplace where the agents have a limited level of collaboration. Another recent paper that studies the equilibrium characterization of a competitive marketplace is Allon et al. (2012). It studies different involvements of the moderating firm in a service marketplace supposing a fixed skill-mix structure. Specifically, the moderating firm can introduce operational tools which provide an efficient match between customers and providers. Moreover, the firms can provide strategic tools which allow communication and collaboration among the agents. Allon et al. (2012) concludes that the moderating firm should compliment its operational tools by creating communication opportunities among providers. Our paper, on the other hand, explores the effects of different skill-mix structures on the moderating firm's revenue, supposing the firm offers both efficient matching between customers and providers and communication among providers.

The research on marketplaces may also be viewed as related to the literature on labor markets that studies the wage dynamics (See Burdett and Mortensen (1998), Manning (2004), and Michaelides (2010)). In this paper, our focus is on a market for temporary help, which means that the engagement between customers and service providers ends upon the service completion. This stands in contrast to the labor economics literature in which the engagement is assumed to be permanent. Furthermore, the entities governing the labor markets can use intervention tools directly influencing the wage dynamic, such as minimum wage. Unlike the literature in labor markets, the moderating firms we consider have minimal direct power to influence the prices emerging as the equilibrium outcomes. Our paper also differs from the literature on market microstructure. This body of work studies market makers who can set prices and hold inventories of assets in order to stabilize markets (See Garman (1976), Amihud and Mendelson (1980), Ho and Stoll (1983), and a comprehensive survey by Biais et al. (2005)). However, the moderating firms considered in our paper have no direct price-setting power and cannot respond to customers' service requests.

3. Model Formulation

Consider a service marketplace where agents and customers make their decisions in order to maximize their individual utilities. There are two groups of customers, each of which has different needs. We use the term "*class*" to identify the group of customers with similar needs. We refer to one customer class as class-A and the other one as class-B. The size of the population of class $i \in \{A, B\}$ is Λ_i, and we assume customers are small relative to the size of the market. This forms the "potential demand" for the marketplace. Each customer decides whether to join the marketplace or not. Customers who join the marketplace form the "effective demand" for the marketplace. If a customer decides not to join the system, she requests the service from an outside option which generates a utility of $\underline{u}$. If she joins the system, she decides who would process her job. The exact nature of this decision depends on the specific structure of the marketplace, which is decided upfront by the moderating firm. We shall elaborate on the choices of customers while we discuss the role of the moderating firm. When the service of a class $i \in \{A, B\}$ customer is successfully completed, she pays the price of the service and earns a reward which depends on the skills of the agent serving her. If the demand from class i exceeds the available service capacity for this class, each class i customer is equally likely to be served. Customers decide whether to request service or not and by whom to be served according to their expected utility. The expected utility of a customer is based on the reward, the price, and the likelihood of being served.

The above summarizes the demand arriving to the marketplace. Next, we discuss the capacity provision in the marketplace. The size of the population of *candidate* agents is k, and agents are small relative to the size of the market. Each agent can serve τ customers but is endowed with different processing skills. Particularly, the value that an agent's service generates for a class $i \in \{A, B\}$ customer is S_i. S_A and S_B are random variables with a joint probability density function $f_{A,B}(\cdot, \cdot)$ on the support $[0, \bar{R}_A] \times [0, \bar{R}_B]$. We refer to (S_A, S_B) as the *agent skills* and $f_{A,B}(\cdot, \cdot)$ as the *skill distribution*. The skills are not observable but the moderating firm can verify whether agent skills are above a threshold through a skill screening process. The decisions of an agent are to set a price for his service and choose the customer class to serve among the classes he is eligible to serve; each agent makes these decisions independently in order to maximize his expected revenue.

The expected revenue of an agent depends on the price he charges and his demand volume. We normalized the operating cost of the agents to zero for notational convenience.

We refer to the ratio $\Lambda_i/(\tau k)$ as the demand-supply ratio of class $i \in \{A, B\}$ and denote it by $\rho_i > 0$. The demand-supply ratio is a first order measure for the mismatch between aggregate demand and the total processing capacity for each class. In this paper, we assume the service capacity of each agent is the same for both classes. Our main results would continue to hold if we relaxed this assumption.

4. The Roles of the Moderating Firm

The essential role of the moderating firm in a large-scale marketplace is to construct the infrastructure for the interaction between players. This is crucial because all players have to be equipped with the necessary information, such as prices to make their decisions, and individual players cannot gather this information on their own. There are also other ways for moderating firms to be involved in a marketplace. For instance, moderating firms can provide mechanisms which improve the operational performance of the whole system by efficiently matching customers and agents. They may also complement their operational tools with strategic tools which enable communication among agents. Furthermore, because agents' skills are not observable to the customers, moderating firms may provide customers with further information about the candidate agents by screening agents' abilities. In this section, our goal is to build a model where we capture these different roles of the moderating firms. To this end, we first introduce a screening mechanism which consists of skill tests determining whether a candidate agent is eligible to serve customers or not. Next, we provide a detailed description of the interaction between customers and agents in a marketplace when operational inefficiencies are minimized and agents are allowed to communicate. Using this model, we will study the moderating firm's skill and capacity management problem with the objective of revenue maximization in Section 6. Note that the firm's only source of revenue is its predetermined share of the total revenue generated in the marketplace. Therefore, the moderating firm uses its screening mechanism to maximize the total revenue in the system.

4.1. Setting up the Skill-Mix

As we mentioned in the introduction, moderating firms can obtain more information about the abilities of candidate agents through a screening process. We model this by assuming that the moderating firm can run two skill tests on each candidate agent, say Exam A and Exam B, in order to screen his abilities. In particular, in Exam $i \in \{A, B\}$, the firm picks a threshold level ω_i (a measure for comprehensiveness) and tests whether the value that an agent's service generates for class i is above ω_i. We refer to these thresholds as the *passing levels*.

The firm publicly announces the results of the tests, and a candidate agent will be *eligible* to serve customers if he passes at least one exam. Candidate agents who pass both exams will be eligible to serve both classes of customers. We refer to these types of agents as *flexible* agents and denote the fraction of flexible agents by α_F. Since a flexible agent can serve both classes, he makes a service decision by choosing which customer class he serves in addition to his pricing decision. On the other hand, an agent who passes only Exam $i \in \{A, B\}$, will be eligible to serve only class i, and thus his only decision will be to set a price for his service. We refer to these types of agents as *dedicated* agents and denote the fraction of dedicated agents for class $i \in \{A, B\}$ by α_i.

Once the firm chooses a pair of passing levels (ω_A, ω_B), the fraction of flexible and dedicated agents in marketplaces with a large number of agents can be approximated as follows:

$$\alpha_F(\omega_A, \omega_B) \simeq \int_{\omega_A}^{\bar{R}_A} ds_A \int_{\omega_B}^{\bar{R}_B} f_{A,B}(s_A, s_B) ds_B \textit{ and } \alpha_i(\omega_A, \omega_B) \simeq \int_{\omega_i}^{\bar{R}_i} ds_i \int_{0}^{\omega_j} f_{A,B}(s_i, s_j) ds_j \tag{1}$$

for all $i, j \in \{A, B\}$ with $j \neq i$. Throughout the paper, we use the term *skill-mix structure* to denote the portfolio of different agent types in the marketplace. For instance, a marketplace may consist of all three types of agents: flexible agents, and dedicated agents for each class. We refer to such a skill-mix structure as "*M-Network*". Moreover, there may be only flexible agents in a marketplace. This skill-mix structure will be referred to as "*V-Network*". In addition to these two, there may be other skill-mix structures such as "*N-Network*" and "*I-Network*". We illustrate these different structures in Figure 1. The moderating firm can set up various skill-mix structures by changing the passing levels ω_A and ω_B.

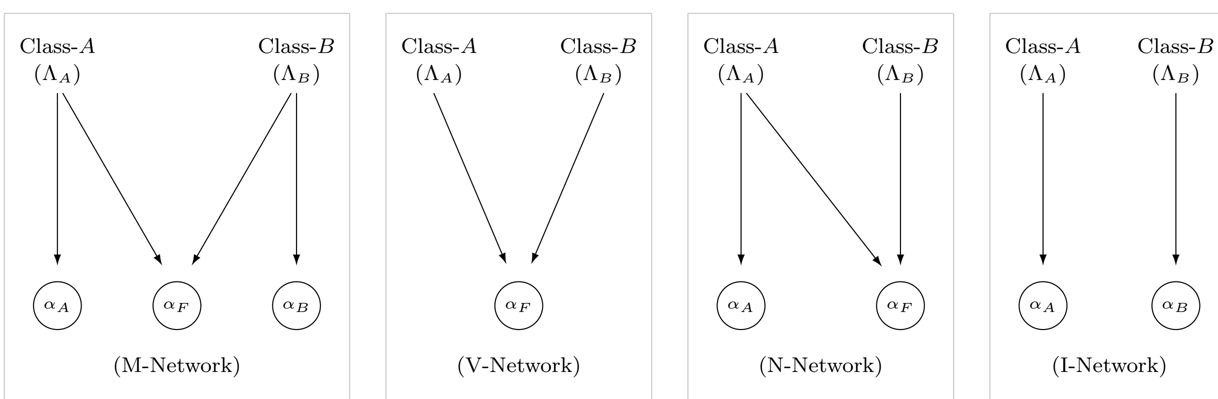


Figure 1 Different skill-mix structures that can be set up by the moderating firm.

In addition to changing the skill-mix structure, passing levels impact the expected reward that a customer earns upon her service completion. For example, when both passing levels are set to zero, a class $i \in \{A, B\}$ customer expects to earn the average value that agents generate for her class, which can be approximated by $\mathbf{E}[S_i]$. However if passing levels are positive, customers can update their expected reward by knowing that the skills of eligible agents are above certain thresholds. More specifically, for any passing levels (ω_A, ω_B) and $i, j \in \{A, B\}$ with $j \neq i$, the expected reward that a class i customer earns from a dedicated agent becomes $\mathbf{E}[S_i | S_i \geq \omega_i, S_j < \omega_j]$ and can be written as follows:

$$R_i(\omega_A, \omega_B) = \Big(\int_{\omega_i}^{\bar{R}_i} ds_i \int_0^{\omega_j} s_i f_{A,B}(s_i, s_j) ds_j \Big) / \alpha_i(\omega_A, \omega_B) \tag{2}$$

Likewise, a class i customer expects to earn $R_{iF}(\omega_A, \omega_B)$ from a flexible agent given passing levels (ω_A, ω_B), where

$$R_{iF}(\omega_A, \omega_B) = \Big(\int_{\omega_i}^{\bar{R}_i} ds_i \int_{\omega_j}^{\bar{R}_j} s_i f_{A,B}(s_i, s_i) ds_j \Big) / \alpha_F(\omega_A, \omega_B). \tag{3}$$

4.2. Matching Demand and Supply

In addition to setting up the skill-mix, the moderating firm provides a mechanism that efficiently matches customers and agents. This mechanism aims at reducing inefficiency due to the possibility that an individual customer may not find an idle agent on her own while there are available agents who can serve her. For instance, upwork.com achieves this goal by allowing customers to post their needs and allowing service providers to apply to these postings. When a customer posts a job at upwork.com, agents that are willing to serve this customer apply to the posting. Among the applicants charging less than what the customer wants to pay, the customer will favor agents charging the lowest price. When there are not any applications, the customer will leave the web-site without being served. The main driver of the operational efficiency in this setting is the fact that customers no longer need to specify an agent upon their arrival. The job posting mechanism allows customers to postpone their service request decisions until they have enough information about the availability of the providers.

Note that the expected utility of a customer will depend on both the price and the type of the agent who serves her because each agent type may provide a different expected reward. To account for that, we define the "*net reward*" of a class $i \in \{A, B\}$ customer from a dedicated agent and a flexible agent charging p as $R_i(\omega_A, \omega_B) - p$ and $R_{iF}(\omega_A, \omega_B) - p$ for any given pair of thresholds (ω_A, ω_B), respectively. Then, we model the efficiency improvement in the system by considering the marketplace as a combination of interrelated agent pools where the agents announcing the same net reward and serving the same customer class are virtually grouped together, regardless of their types.

Once each agent announces a price per customer to be served, we can construct a resulting net reward vector $(r_{i_n})_{n=1}^{N_i}$ for class $i \in \{A, B\}$ where N_i is the number of different net rewards announced by the agents serving class i. We refer to the agents announcing the net reward r_{i_n} as sub-pool i_n and denote the fraction of agents in the sub-pool i_n by y_{i_n}. Hence, the vectors $(\mathbf{r_A}, \mathbf{y_A}) \equiv (r_{A_n}, y_{A_n})_{n=1}^{N_A}$ and $(\mathbf{r_B}, \mathbf{y_B}) \equiv (r_{B_n}, y_{B_n})_{n=1}^{N_B}$ summarize the strategy of all agents.

Under the mechanism provided by the moderating firm, we model customer decision making and experience as follows: If there are different net rewards announced by the agents serving class $i \in \{A, B\}$, i.e. $N_i > 1$, a class i customer chooses a sub-pool from which she requests the service. We refer to the net reward offered by this sub-pool as the "*preferred net reward*." We denote the fraction of customers requesting service from sub-pool i_n by D_{i_n} and summarize the decisions of the class i customers by the vector $\mathbf{D_i} \equiv (D_{i_n})_{n=1}^{N_i}$ for all $i \in \{A, B\}$. As we model the marketplace as a combination of interrelated agent pools, the operations of each sub-pool depend on the operations of the other sub-pools. To be more specific, each sub-pool may handle customers from other sub-pools while some of the other sub-pools are serving its customers. Therefore, each customer requesting service may experience three possible outcomes: 1) being assigned to a sub-pool offering a net reward higher than her preferred net reward, 2) being assigned to her preferred sub-pool, and 3) being rejected for service. We assume that customers will have priority in their preferred sub-pools over the customers who may come from other sub-pools. Then, we can write the expected utility of a class $i \in \{A, B\}$ customer choosing the sub-pool i_ℓ with $\ell \in \{1, \ldots, N_i\}$ as[1]

$$U_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i}) = \sum_{n=1}^{N_i} P_{i_{\ell n}}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i}) r_{i_n}, \tag{4}$$

where $P_{i_{\ell n}}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i})$ denotes the probability that a customer choosing the sub-pool i_ℓ is served by the sub-pool i_n for all $n \in \{1, \ldots, N_i\}$. We want to note that for any sub-pool i_ℓ, $P_{i_{\ell n}} = 0$ for any n such that $r_{i_n} < r_{i_\ell}$ since customer choosing sub-pool i_ℓ cannot be served by a sub-pool offering a net reward less than their preferred net reward, r_{i_ℓ}. Furthermore, given the passing levels (ω_A, ω_B), the revenue of a dedicated agent in the sub-pool i_ℓ is $\left[R_i(\omega_A, \omega_B) - r_{i_\ell}\right] \sigma_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i})\tau$ where $\sigma_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i})$ is the utilization of each agent, and thus $\sigma_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i})\tau$ is the average demand in sub-pool i_ℓ. Similarly, the revenue of a flexible agent in the sub-pool i_ℓ is $\left[R_{iF}(\omega_A, \omega_B) - r_{i_\ell}\right] \sigma_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i})\tau$.

One can notice that both $P_{i_{\ell n}}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i})$ and $\sigma_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i})$ depend on the strategies of agents and customers and may take quite complicated forms. In order to clarify the definitions of these functions, we illustrate the operations of an example sub-pool, say sub-pool i_ℓ for some $\ell \in \{1, \ldots, N_i\}$ and $i \in \{A, B\}$ in Figure 2.

[1] Here, we assume that a customer with a preferred net reward r_{i_ℓ} receives a net reward of r_{i_m} when she is served by sub-pool i_m for any $m \neq \ell$. Our key findings would continue to hold, despite additional analytical complexity, if this assumption was relaxed.

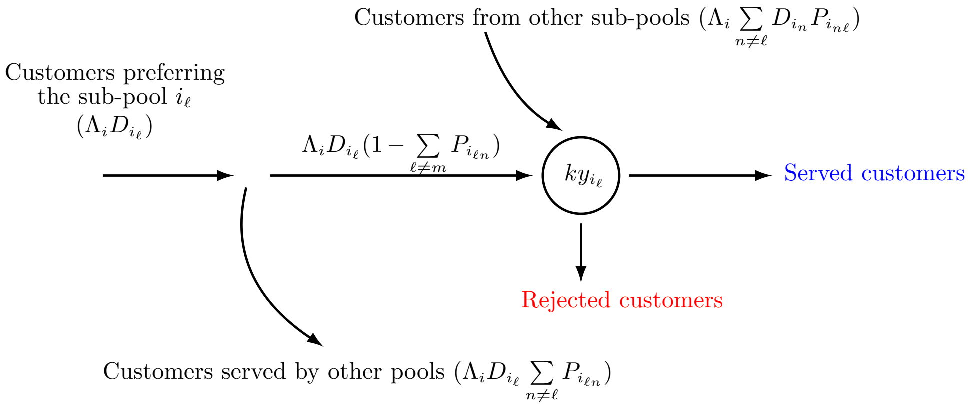


Figure 2 **An example sub-pool.**

Note that customers preferring sub-pool i_ℓ have priority over the customers who come from other sub-pools. Therefore, the number of customers who prefer and are served by sub-pool i_ℓ is the minimum of demand, which is $\Lambda_i D_{i_\ell}(1-\sum_{n\neq\ell} P_{i_{\ell n}})$, and the available service capacity, which is $k\tau y_{i_\ell}$. Then, since customers are assigned to agents randomly, we can write the probability with which customers choosing the sub-pool i_ℓ are served by the sub-pool i_ℓ as:

$$P_{i_{\ell\ell}}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i}) = \min\Big\{\frac{k\tau y_{i_\ell}}{\Lambda_i D_{i_\ell}(1-\sum_{n\neq\ell} P_{i_{\ell n}})}, 1\Big\}.$$

Similarly, the number of customers who are served by sub-pool i_ℓ is the minimum of total demand, which includes both the customers preferring the sub-pool i_ℓ and customers coming from other sub-pools, and the available service capacity, which is $k\tau y_{i_\ell}$. Then, we can write the utilization of each agent in the sub-pool i_ℓ as:

$$\sigma_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i}) = \min\Big\{\frac{\Lambda_i\big[D_{i_\ell}(1-\sum_{n\neq\ell} P_{i_{\ell n}}) + \sum_{n\neq\ell} D_{i_n} P_{i_{n\ell}})\big]}{k\tau y_{i_\ell}}, 1\Big\}.$$

After describing the operational tool provided by the moderating firm, we now discuss the moderating firm's strategic tool, which changes the nature of the competition among agents. In a marketplace such as upwork.com, service providers are offered discussion boards where they are allowed to exchange information. Moreover, the firm supports the creation of affiliation groups, which are self-enforcing entities. Motivated by these examples, we assume that the moderating firm allows agents to make non-binding communication prior to making their decisions, so that the players can discuss their strategies but are not allowed to make binding commitments. The economics literature suggests that the stability of any equilibrium outcome can be threatened by potential deviations formed by coalitions, even in noncooperative games, due to pre-play communications (See Ray (1996) and Moreno and Wooders (1996)). In other words, players can try to self-coordinate their actions in a mutually beneficial way when the moderating firm allows them to communicate among themselves, despite the fact that each agent selfishly maximizes his own utility. As it is discussed in Allon et al. (2012), this can be modeled by an equilibrium concept which allows several agents to deviate together instead of deviating individually. It is crucial to consider group deviations in our equilibrium concept because customers might fail to notice individual agent deviations due to the fact that agents are small relative to the size of the market. Furthermore, self-coordination of agents is restricted because the marketplaces we consider tend to be large. We denote the largest fraction of agents that is allowed to deviate together by $\delta \leq 1$.

In this section, we outline the different roles of the moderating firm and discuss how these roles affect the structure of the marketplace. Next, we model the strategic interaction between the agents and the customers as a sequential-move game given the setup of Section 3, along with the above mentioned roles of the moderating firm. We also characterize the equilibrium outcome of a special marketplace structure.

5. The Game Between Customers and Agents

In this section, we formally set up the two-stage game between the agents and the customers based on the model introduced in Section 3 and the roles of the moderating firm mentioned in Section 4. As the first step of the strategic interaction between agents and customers, agents make and announce their service and pricing decisions. Then, in the second stage, each arriving customer observes these decisions and decides whether to request service or not. Customers also specify the sub-pool from which they request service. We suppose that agents make the first move because agents have permanent profile pages, where they post pricing information for their services, in real online marketplaces. Moreover, posting a price is neither a requirement nor a binding decision for customers in job postings. Therefore, customers usually do not announce any price in job postings because they prefer to wait and see the agents' availability and prices.

We refer to the equilibrium among customers in the second stage as *Customer Equilibrium* and the equilibrium of the whole game as *Market Equilibrium*. In the following subsections, we introduce the definitions of these equilibrium concepts and characterize their outcomes. Since we study a two-stage game, we start with the equilibrium among customers given the strategy of agents.

5.1. Customer Equilibrium

As we mentioned before, customers make their service request in order to maximize their expected utility. Therefore, in the equilibrium of the second stage, a customer from class $i \in \{A, B\}$, chooses a sub-pool only if the expected utility she obtains from this sub-pool (weakly) dominates her utility from any other sub-pool. The first condition of the *Customer Equilibrium* captures this requirement as we formally define as follows:

DEFINITION 1 (CUSTOMERS EQUILIBRIUM). Given any $(\mathbf{r_i}, \mathbf{y_i})$ for $i \in \{A, B\}$, we say that $\mathbf{D_i} \equiv (D_{i_n})_{n=1}^{N_i}$ is a *Customers Equilibrium* if the following conditions are satisfied:

1. For any ℓ with $D_{i_\ell} > 0$ and for all $n \leq N_i$, we have that $U_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i}) \geq U_{i_n}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i}) \geq \underline{u}$.
2. If $U_{i_n}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i}) > \underline{u}$ for some $n \leq N_i$, then $\sum_{n=1}^{N_i} D_{i_n} = 1$.
3. If $U_{i_n}(\mathbf{r_i}, \mathbf{y_i}, \mathbf{D_i}) = \underline{u}$ and $\rho D_{i_n} < y_{i_n}$ for some $n \leq N_i$, then $\sum_{n=1}^{N_i} D_{i_n} = 1$.

In addition to stating that customers maximize their expected utility, in its second condition, the *Customer Equilibrium* ensures that all customers join the system if it is possible to earn an expected utility that is strictly greater than their outside option, which generates a utility of $\underline{u}$. Finally, the third condition states that all customers join the system when there is an under-utilized sub-pool if their expected utility is equal to their outside option. The last condition essentially breaks ties for a customer who is indifferent between joining the system and leaving immediately in favor of joining. A *Customer Equilibrium* always exits according to Rath (1992) since the utility functions are continuous. However, the uniqueness of a *Customer Equilibrium* is not guaranteed for a given strategy of agents. Although there may be multiple equilibria for a given strategy of agents, Proposition 1 proves that the equilibrium utilization of an agent can be characterized irrespective of the multiplicity in *Customer Equilibrium.*

PROPOSITION 1. *Given any $(\mathbf{r_i}, \mathbf{y_i})$ such that $r_{i_n} > r_{i_{n+1}}$ for all $n \in \{1, \ldots, N_i\}$ and $i \in \{A, B\}$, let $\sigma^{ce}_{i_\ell}(\mathbf{r_i}, \mathbf{y_i})$ be the utilization of an agent in sub-pool i_ℓ in a Customer Equilibrium for all $\ell \in \{1, \ldots, N_i\}$. Then, we have that*

$$\sigma^{ce}_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}) = \begin{cases} 1 & \text{if } \rho_i > \rho^0_\ell + y_{i_\ell}, \\ (\rho_i - \rho^0_n)/y_{i_n} & \text{if } \rho \in [\rho^0_\ell, \rho^0_\ell + y_{i_\ell}], \\ 0 & \text{if } \rho_i < \rho^0_\ell. \end{cases}$$

Here, we let $\rho^0_\ell = \big(\sum_{n=1}^{\ell-1} r_{i_n} y_{i_n}\big)/r_{i_\ell}$, and $\rho^0_1 = 0$.

The above proposition shows that agents in a sub-pool can earn a strictly positive revenue only if the demand from the customer class they serve is greater than a critical demand level. The critical demand level for each sub-pool depends only on the decisions of the sub-pools offering a net reward that is higher than what the given sub-pool offers. Furthermore, Proposition 1 shows that a sub-pool can be fully utilized when demand rate exceeds the critical demand level for this sub-pool by at least its capacity. In Proposition 1, we focus only on the utilization of the agents in the customer equilibrium. We can also characterize the fraction of customers choosing each sub-pool explicitly (See Proposition 3 in Appendix A.2). The main driver of the result in Proposition 1 is that customers are either always or never served by sub-pools other than their preferred one in equilibrium. This establishes that the interdependency between the sub-pools announcing different net rewards is minimal in a *Customer Equilibrium*. In fact, the marketplace operates "almost like" the combination of independent sub-pools.

Once we characterize the equilibrium among customers and obtain the revenues of agents in this equilibrium, we now focus on the first stage of the game.

5.2. Market Equilibrium

The *Customer Equilibrium* we study in the previous subsection derives agent revenues when we fix the service and pricing decisions of the agents. Using this result, we now study the equilibrium outcome of the whole game, which will be referred to as the *Market Equilibrium*. To this end, we need to find service and pricing decisions from which agents have no incentive to deviate in the first stage. Agents can deviate by either joining an existing sub-pool or announcing a new price. Furthermore, a limited fraction of agents are allowed to deviate together since the moderating firm enables communication among agents. Therefore, an equilibrium in the first stage should be immune to any of these two types of deviations formed by at most δ fraction of agents. In the large-scale marketplace we study, it is possible that a small group of agents can find profitable deviation from every price in some cases. However, the gains from some deviations may be arbitrarily small. As in Allon et al. (2012), we ignore deviations which result in small gains by employing a somewhat weaker notion of equilibrium. This equilibrium notion allows us to characterize the market outcome even when a Nash equilibrium does not exist. To be more specific, we focus on an equilibrium concept which requires immunity against only deviations that improve the revenue of an agent by at least $\epsilon > 0$ as formally stated in Definition 2. We refer to ϵ as the level of equilibrium approximation and suppose it is arbitrarily close to zero. To ease notation, we denote $R_i(\omega_A, \omega_B)$ and $R_{iF}(\omega_A, \omega_B)$ by R_i and R_{iF}, respectively, for any given passing levels (ω_A, ω_B) and any $i \in \{A, B\}$. We also let $\mathbf{e_i}^x \equiv (e^x_{i_n})_{n=1}^{N_i}$ denote the N_i-dimensional vector with a 1 in the x^{th} coordinate and 0 elsewhere.

Definition 2 (Market Equilibrium). Let $(\mathbf{r_i}, \mathbf{y_i}) \equiv (r_{i_n}, y_{i_n})_{n=1}^{N_i}$ summarize the strategy of all agents in the marketplace for any $i \in \{A, B\}$. Then, $(\mathbf{r_i}, \mathbf{y_i})$ is a (ϵ, δ)-*Market Equilibrium ((ϵ, δ)-ME)* if the following conditions are satisfied.

1. For any $\ell \le N_i$, $m \le N_i$, $0 < d \le \min\{y_{i_\ell}, \delta\}$, and $i \in \{A, B\}$, we have that

$$[R_i - r_\ell]\sigma^{ce}_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}) \ge [R_i - r_m]\sigma^{ce}_{i_m}(\mathbf{r_i}, \mathbf{y'_i}) - \epsilon \tag{5}$$

$$[R_{iF} - r_\ell]\sigma^{ce}_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}) \ge [R_{iF} - r_m]\sigma^{ce}_{i_m}(\mathbf{r_i}, \mathbf{y'_i}) - \epsilon, \ \textit{where } \mathbf{y'_i} = \mathbf{y_i} - d\mathbf{e_i}^\ell + d\mathbf{e_i}^m. \tag{6}$$

2. For any $\ell \le N_i$, $0 < d \le \min\{y_{i_\ell}, \delta\}$, $i \in \{A, B\}$, and $r' \ne r_{i_n}$ for all $n \le N_i$, we have that

$$[R_i - r_\ell]\sigma^{ce}_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}) \ge [R_i - r']\sigma^{ce}_{i_{N_i+1}}(\mathbf{r'_i}, \mathbf{y'_i}) - \epsilon, \tag{7}$$

$$[R_{iF} - r_\ell]\sigma^{ce}_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}) \ge [R_{iF} - r']\sigma^{ce}_{i_{N_i+1}}(\mathbf{r'_i}, \mathbf{y'_i}) - \epsilon, \ \textit{where } \mathbf{r'_i} = (\mathbf{r_i}, r'), \ \textit{and } \mathbf{y'_i} = (\mathbf{y_i}, d). \tag{8}$$

3. For any $\ell \leq N_i$, $m \leq N_j$, $0 < d \leq \min\{y_{i_\ell}, \delta\}$, $j \in \{A, B\} \setminus i$, we have that

$$[R_{iF} - r_\ell]\sigma^{ce}_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}) \geq [R_{jF} - r_m]\sigma^{ce}_{j_m}(\mathbf{r_j}, \mathbf{y'_j}) - \epsilon, \; \textit{where } \mathbf{y'_j} = \mathbf{y_j} + d\mathbf{e_j}^m. \quad (9)$$

4. For any $\ell \leq N_i$, $0 < d \leq \min\{y_{i_\ell}, \delta\}$, $j \in \{A, B\} \setminus i$, $r' \neq r_{j_n}$ for all $n \leq N_j$, we have that

$$[R_{iF} - r_\ell]\sigma^{ce}_{i_\ell}(\mathbf{r_i}, \mathbf{y_i}) \geq [R_{jF} - r']\sigma^{ce}_{j_{N_j+1}}(\mathbf{r'_j}, \mathbf{y'_j}) - \epsilon, \; \textit{where } \mathbf{r'_j} = (\mathbf{r_j}, r'), \; \textit{and } \mathbf{y'_j} = (\mathbf{y}, d). \quad (10)$$

Moreover, $(\mathbf{r_i}, \mathbf{y_i})$ is a *Market Equilibrium* if there exists a sequence $(\mathbf{r_i}^k, \mathbf{y_i}^k)$ such that $(\mathbf{r_i}^k, \mathbf{y_i}^k)$ is a (ϵ^k, δ^k)-*ME* where $\epsilon^k \to 0$, $\delta^k \to 0$, and $(\mathbf{r_i}^k, \mathbf{y_i}^k) \to (\mathbf{r_i}, \mathbf{y_i})$ for all $n \leq N_i$ as $k \to \infty$.

(5) and (7) in the (ϵ, δ)-*ME* definition state that dedicated agents have no incentive to deviate. Note that dedicated agents cannot change the customer class they serve. Therefore, (ϵ, δ)-*ME* accounts for two possible deviations for dedicated agents: joining an existing sub-pool or creating a new one. On the other hand, flexible agents have the option of changing the customer class they serve. Thus, (ϵ, δ)-*ME* ensures that flexible agents cannot improve their revenues whether they change the customer class they serve or not. In particular, (6) and (8) focus on flexible agent deviations when they keep the customer class they serve the same, whereas (9) and (10) consider deviations where flexible agents change the class they serve. Finally, we conclude that a strategy profile is a *Market Equilibrium* if it is the limit of a sequence of (ϵ, δ)-*ME* as ϵ and δ become arbitrarily small. By doing so, we lessen the role of ϵ and δ in our equilibrium definition.

It is quite tedious to study the *Market Equilibrium* of the whole marketplace in detail because the moderating firm can create many different *skill-mix structures* as discussed in Section 4.1. Hence, we next characterize the equilibrium outcome in a marketplace with a special market structure, namely one customer class and two types of agents. This structure constitutes the building block of a marketplace with two classes. Carrying out our analysis in this building block model is a fundamental step towards finding the equilibrium outcome of the whole marketplace. It also allows us to discuss the intuition behind our results in a more clear way.

5.3. A Special Marketplace Structure

As we mentioned before, the moderating firm may set up the skill-mix structure in the marketplace (i.e., the capacity of dedicated and flexible agents) by changing the passing levels in each exam. Figure 1 illustrates the possible skill-mix structures that can arise based on the firm's skill screening decision. We observe that *Inverted V-Network*, where there is only one customer class and two types of agents, is the key to characterizing the equilibrium outcome in any skill-mix structure. For instance, once the agents make their service decisions in an *M-Network*, we need to analyze two separate *Inverted V-Networks*, in each of which one customer class can be served by both dedicated and flexible agents. Similarly, we can obtain the equilibrium outcome in other skill-mix structures by analyzing two separate *Inverted V-Networks*. Therefore, in this subsection, we study the equilibrium outcome of this special market structure: the *Inverted V-Network*.

We consider a marketplace where there is only one class of customers, but two types of agents, say dedicated and flexible, who can serve these customers. We assume that the size of the customer population is Λ. Thus, the demand-supply ratio is $\rho \equiv \Lambda/(k\tau)$. As we have two types of agents, we let α_D be the fraction of dedicated agents, and α_F be the fraction of flexible agents. Customers earn a reward of R_D when their service is completed by a dedicated agent, and they earn R_F when a flexible agent serves them. We suppose $R_D \geq R_F$ for ease of explanation. Our results would only need relabeling if $R_F \geq R_D$. We keep all other assumptions we made in Section 3, and we will use the equilibrium concepts introduced in Sections 5.1 and 5.2. We refer to the marketplace as *buyer's market* if $\rho < \alpha_D + \alpha_F$ and *seller's market* otherwise.

Theorem 1 formally presents the Market Equilibrium in the special market structure. An important implication of our equilibrium result is that the revenue of the flexible agents cannot exceed their operating cost, which is normalized to zero, in a buyer's market. The main driver of this result is that the dedicated agents can price the flexible agents out of the market when $\rho < \alpha_D + \alpha_F$ since customers earn a higher reward from the dedicated agents. On the other hand, both groups of agents can agree to charge the customer reward they generate minus the outside option in a seller's market. In such an equilibrium, agents charge their highest prices and are fully utilized, and the rate of customers requesting service is equal to the total capacity, $\alpha_D + \alpha_F$. In other words, the equilibrium behavior of the agents ensures that the demand perfectly matches supply in the system while extracting all of the customer surplus. The intuition behind these high prices is the following: The customer demand exceeds the total service capacity when agents charge lower than the highest price that agents can charge. Therefore, customers face a strictly positive probability of being rejected for service. This allows agents to increase their prices because customers will be willing to pay a premium for a higher chance of getting served. As long as the price increase is small enough, deviating agents improve their revenues.

THEOREM 1. *Let V_D^{sm} and V_F^{sm} be the revenue of a dedicated and a flexible agent, respectively, in a Market Equilibrium of a marketplace with one customer class and two agent pool.*

1. *If $\rho < \alpha_D$, then we have that $V_D^{sm} = V_F^{sm} = 0$.*
2. *If $\rho = \alpha_D$, then we have that $V_D^{sm} \leq (R_D - R_F)\tau$, and $V_F^{sm} = 0$.*
3. *If $\alpha_D < \rho < \alpha_D + \alpha_F$, then we have that $V_D^{sm} = (R_D - R_F)\tau$, and $V_F^{sm} = 0$.*
4. *If $\rho = \alpha_D + \alpha_F$, then we have that $V_D^{sm} = V_F^{sm} + (R_D - R_F)\tau$, and $V_F^{sm} \leq (R_F - \underline{u})\tau$.*
5. *If $\rho > \alpha_D + \alpha_F$, then we have that $V_D^{sm} = (R_D - \underline{u})\tau$, and $V_F^{sm} = (R_F - \underline{u})\tau$.*

Two of the simplifying assumptions we impose are that (i) all customers and agents are present in the marketplace at the same time, and (ii) the time required to serve customers is fixed and known to all players. One may envision a more dynamic and stochastic interaction between customers and agents, specifically by a model where the service requests of customers arrive according to a stochastic process and service times are random. We confirmed that our findings in Theorem 1 carry over to this richer and more complicated model. The equilibrium outcomes of the above building block marketplace structure is the key driver to solving the moderating firm's problem. Thus, our key managerial insights into the skill and capacity management decisions of the firm also continue to hold in the more dynamic model.

After characterizing the equilibrium outcome in our building block model, we next turn our attention to the moderating firm's skill-mix and capacity decisions.

6. The Moderating Firm's Problem

In the previous section, we analyze a model where the skill-mix of the marketplace is given. In this section, we study the firm's problem of finding the best skill-mix structure that maximizes its revenue. The firm's revenue is a predetermined share of the total revenue in the marketplace. Thus, the firm has to maximize the total revenue in the marketplace by choosing the appropriate passing levels (ω_A, ω_B). Once the firm chooses q skill-mix structure via the skill tests, the flexible agents make their service decision. Based on the decisions of the flexible agents, the firm's business with each class operates as the market structure we study in Section 5.3. Then, we can can calculate the revenue generated in the marketplace using Theorem 1. Letting $\overline{V}_{iF}(\omega_A, \omega_B)$ and $\overline{V}_{iD}(\omega_A, \omega_B)$ be the equilibrium revenues of a flexible and dedicated agent serving class $i \in \{A, B\}$, respectively[2], the total revenue of the marketplace for any passing levels (ω_A, ω_B) is

$$\Pi(\omega_A, \omega_B) = k\big[\alpha_F(\omega_A, \omega_B)\overline{V}_{iF}(\omega_A, \omega_B) \sum_{i \in \{A,B\}} \alpha_i(\omega_A, \omega_B)\overline{V}_{iD}(\omega_A, \omega_B)\big].$$

The main focus of this paper is to gain a better understanding of the effectiveness of skill screening as a revenue management tool. As a first step toward this goal, we study the firm's problem under the following three cases: (i) *Benchmark*, where $\omega_A = \omega_B = 0$, (ii) *One-Test*, where either $\omega_A = 0$ or $\omega_B = 0$ but not both, and (iii) *Two-Tests*, where both $\omega_A > 0$ and $\omega_B > 0$. Considering these three cases separately allows us to analyze the revenue improvements in the marketplace after each additional test. We, then, find the optimal number of exams when the skill screening is costly.

[2] Theorem 1 shows that there might be multiple equilibria if the demand rate is equal to the capacity. For mathematical convenience, we focus on the equilibrium generating the highest possible revenue among all possible equilibria. The firm can sustain an equilibrium where the revenue arbitrarily close to the highest level among the multiple equilibria by perturbing the passing levels (ω_A, ω_B) slightly and creating a seller's market with a unique equilibrium.

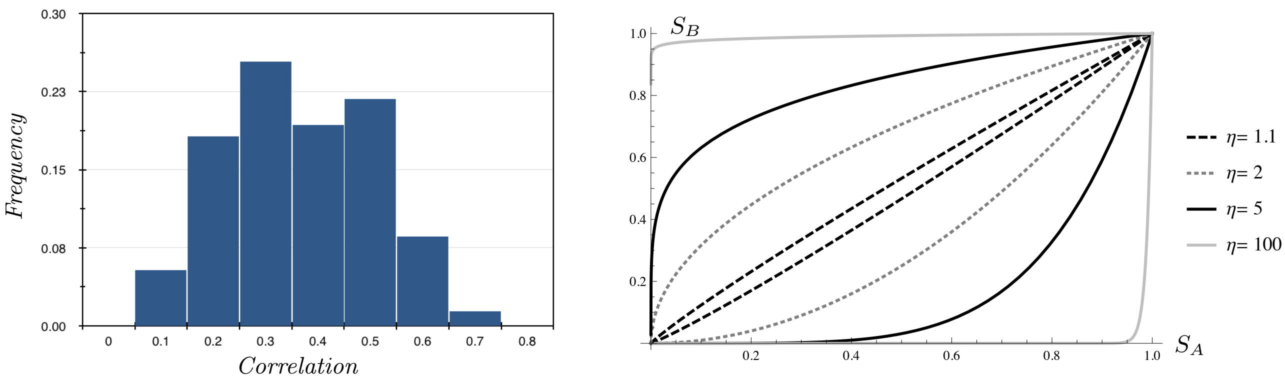
A major technical challenge in finding the firm's optimal decisions is that the revenue function may have different functional forms in different regions of passing level space $[0,\bar{R}_A]\times[0,\bar{R}_B]$ because equilibrium revenues of the agents can change significantly even for slight adjustments in the passing levels. Finding the optimal passing levels, then, requires comparing all these different functional forms, which becomes analytically intractable when the skills, S_A and S_B, follow a general joint probability distribution. For tractability of the firm's problem, we need to impose some assumptions on the skill distributions.

The marketplaces we study, in general, attract service providers who are endowed with similar skills. For example, upwork.com has many agents with various programming skills. Therefore, it is not unrealistic to suppose that agent skills are positively correlated. We also observe that the agent skills are mostly positively correlated from the data we collected from upwork.com. In Table 1, we present the correlation between a pair of exams that are taken by at least 500 unique service providers.

	1	2	3	4	5	6	7
1: U.S. Eng. Skills	1.000						
2: English Spelling	0.335	1.000					
3: Office Skills	0.568	0.286	1.000				
4: Email Etiquette	0.559	0.366	0.576	1.000			
5: HTML 4.01	0.398				1.000		
6: PHP5					0.487	1.000	
7: Customer Service	0.536	0.299	0.611	0.601			1.000
8: CSS 2.0					0.563		
9: Search Engine Opt.		0.175					
10: Call Cent. Skills	0.487	0.384		0.612			0.537
11: UK Eng. Skills	0.788	0.344					

Table 1 Correlation between exam pairs that are taken at least 500 times

We also plot the frequency of the correlation coefficients between exam pairs that are taken at least 100 times in Figure 3. As one can see in Figure 3.a, correlation coefficients between exam pairs range between 0.1 (almost independent) and 0.8 (almost perfectly correlated). In order to capture these observations, we focus on a family of joint skill distributions with a shape parameter η. Specifically, we suppose $f_{A,B}(s_A,s_B;\eta)=\frac{\eta+1}{\eta-1}$ for any $0\le S_A^{\eta}\le S_B\le S_A^{1/\eta}\le 1$ and zero otherwise for a given shape parameter $1<\eta<\infty$. We illustrate the support of the agent skill distribution for different values of the shape parameter in Figure 3.b.



(a) (b)

Figure 3 **(a) Frequencies of correlation coefficients in upwork.com data. (b) Support of the $f(s_A, s_B; \eta)$ for various values of the shape parameter η (the area between curves).**

As we formally present in Proposition 2, the correlation coefficient between S_A and S_B, denoted by $Corr(S_A, S_B)$, varies between zero and one. Furthermore, agent skills become independent and identically distributed as η approaches ∞ and perfectly correlated as η approaches one. We also note that the marginal distributions of S_A and S_B are symmetric, so that the average value that an agent's service generates for both classes are the same, i.e., $\mathbf{E}[S_A] = \mathbf{E}[S_B]$. Henceforth, we will refer to these average values as *average skill* and denote it by $\mathbf{E}[S_\eta]$ for any given shape parameter η. We also let $F_\eta(\cdot)$ denote the marginal distribution of each agent skill.

Proposition 2. *For any $1 < \eta < \infty$, we have that $0 < Corr(S_A, S_B) < 1$. Furthermore, $\lim_{\eta\to\infty} Corr(S_A, S_B) = 0$ and $\lim_{\eta\to 1} Corr(S_A, S_B) = 1$.*

In the following subsections, we restrict our attention to a market where the total demand rate exceeds the total service capacity of candidate agents, and normalize the utility of customers from their outside option, $\underline{u}$, to zero. We provide a brief discussion on the firm's optimal decisions when these assumptions are relaxed in the Conclusion. We also suppose $\rho_A \geq \rho_B$ without loss of generality. It is also worth mentioning that our methodology and model allow us to study negatively correlated skills. For instance, we could focus on the skills that satisfy $0 \leq 1 - S_A^{1/\eta} \leq S_B \leq 1 - S_A^{\eta} \leq 1$.

6.1. Benchmark

We start our analysis by considering the *Benchmark* case, where the firm does not offer any skill tests. One can view this case as if passing levels are set to zero, i.e. $(\omega_A, \omega_B) = (0, 0)$. We denote the total revenue of the marketplace in the *Benchmark* case by Π^o.

In the *Benchmark* case, all agents are eligible to serve both customer classes and can choose the customer class they would like to serve. In other words, the skill-mix structure is a *V-Network* with only flexible agents. Furthermore, customers' expected reward is equal to the average skill $\mathbf{E}[S_\eta]$ in the absence of any skill tests.

As we have only flexible agents in the *Benchmark* case, each class of customers face only one pool of agents once the agents make their service decisions. As we show in Theorem 1, agents serving class $i \in \{A, B\}$ can extract all of the customer surplus by charging $\mathbf{E}[S_\eta]$ if their capacity is less the demand from class i and earn zero otherwise. Hence, agents will always prefer to be in a seller's market where the service capacity is scarce. Since the total demand rate is higher than potential service capacity, i.e., $\rho_A + \rho_B \geq 1$, agents can, indeed, sustain an equilibrium where customers from both classes face a seller's market and pay their expected reward. The following theorem formally presents this result and the total revenue of the marketplace under the *Benchmark* case.

THEOREM 2. *The total revenue of the marketplace in the Benchmark case is* $\Pi^o = k\tau \frac{(\eta+1)^2}{(\eta+2)(2\eta+1)}$.

6.2. One Skill Test

After analyzing the *Benchmark* case, we now study the *One-Test* case, where the firm offers only one skill test. As each agent has two skills, the firm has to choose the skill to be tested and the passing level of the test. The firm's objective, in the *One-Test* case, is then to maximize the revenue, $\Pi(\omega_A, \omega_B)$, given the constraints of $\omega_A \omega_B = 0$ and $\omega_A + \omega_B > 0$.

In the *One-Test* case, agents who pass the skill test are eligible to serve both classes while the failing ones can only serve the customers who request service related to the skill that is not tested. In other words, the skill-mix structure is an *N-Network* with a flexible and a dedicated agent pool. Throughout this section, we suppose the firm offers only exam $i \in \{A, B\}$, so that dedicated agents are only eligible to serve class $j \neq i \in \{A, B\}$ customers. We denote the optimal revenue of the marketplace by Π_i^* and the optimal passing level by ω_i^* . We also define the relative improvement in revenue from the *Benchmark* case to the *One-Test* case as $\Pi_i^*/\Pi^o - 1$ and denote it by Δ_i^*.

When the moderating firm sets the passing level as ω, the fraction of dedicated and flexible agents are $F_\eta(\omega)$ and $1 - F_\eta(\omega)$, respectively. Our results in Theorem 1 suggests that the dedicated agents cannot generate any revenue if their service capacity exceeds the demand from class j, i.e. $F_\eta(\omega) > \rho_j$. We show that the firm has to avoid these equilibrium outcomes resulting in zero revenue for the dedicated agents, and thus set $\omega \leq F_\eta^{-1}(\rho_j)$. We also show that the firm prefers flexible agents to serve only class i customers because class i expects a higher reward from flexible agents than class j does. To ensure all flexible agents serve class i, the firm must choose a passing level that is greater than $F_\eta^{-1}(1 - \rho_i)$ according to our results in Theorem 1. These two bounds on ω establish that the firm only needs to consider the interval of $[F_\eta^{-1}(1 - \rho_i), F_\eta^{-1}(\rho_j)]$ while choosing the optimal passing level. We refer to this interval as the *dominating interval.*

In the dominating interval, we show that firm's revenue from each class increases by the service capacity allocated to this class. As the moderating firm cannot increase the service capacity for both classes simultaneously, it trades off between the gains from increasing the capacity for one class and the losses from decreasing the capacity for the other class. It turns out that the gains from increasing the service capacity for class j dominate the firm's losses from class i when the demand from class j is lower than a critical demand level $\overline{\rho}$. Thus, if $\rho_j < \overline{\rho}$, the moderating firm increases the service capacity for class j until the capacity meets the demand from this class. Then, it allocates the rest of the agents to class i. As a result of this skill-mix structure, all of the class j customers obtain service, whereas some customers from class i do not request service since there is not enough capacity to serve the entire class i. Similarly, the firm maximizes its revenue by serving all customers in class i when the demand from class i is lower than the critical level of $1-\overline{\rho}$. Finally, if the demands from both classes exceed the corresponding demand thresholds, the moderating firm maximizes its revenue at a passing level where the gains from increasing the capacity for one class is equal to the losses from decreasing the capacity for the other class. We formally present the above results in the following theorem.

THEOREM 3. *The optimal passing threshold when the firm offers exam $i \in \{A, B\}$ is*

$$\omega_i^* = \begin{cases} F_\eta^{-1}(\rho_j) & \text{if } \rho_j \le \overline{\rho} \\ \overline{\omega}(\eta) & \text{if } 1-\rho_i < \overline{\rho} < \rho_j \\ F_\eta^{-1}(1-\rho_i) & \text{if } 1-\rho_i \le \overline{\rho}, \end{cases}$$

where $\overline{\rho} = F_\eta(\overline{\omega})$ and $\overline{\omega}$ is the unique non-trivial solution for $\omega^{1/\eta} + \omega^\eta - 2\omega = 0$.

The above theorem requires to solve an analytically intractable fixed point problem to obtain an explicit form for the optimal passing level. Thus, it is not possible to use the above theorem to evaluate how much the firm benefits from an additional exam. To obtain insights for the revenue improvements gained in the *One-Test* case, we study the firm's problem under the limiting cases of the skill distribution, namely when the shape parameter η approaches infinity or one.

In our first limiting case, we let the shape parameter η grow to infinity. Proposition 2 shows that the skills of an agent become independently distributed when η approaches infinity. Furthermore, the marginal skill distribution, $F_\eta(\cdot)$, becomes a *Uniform* distribution in this limiting case. As the skill distribution approaches *Uniform*, we show that the firm is indifferent between offering exams A and B. It is interesting that regardless of the exam type, the firm optimally allocates enough capacity to serve all customers from class B (lower demand class), as long as their demand rate is less than the critical level of 1/2. However, the firm always makes some of the potential customers from class A leave without getting the service. When the demand from class B is above 1/2, the firm distributes the service capacity equally, and thus some of the customers from both classes end up not requesting service. We also show that offering an additional exam may improve the firm's revenues as much as 25% compared to the *Benchmark* case.

Our second limiting case assumes the shape parameter η approaches one, and thus the agent skills become perfectly and positively correlated, as shown in Proposition 2. When the skills of an agent are perfectly correlated, the firm's gains from increasing the service capacity for one class turns out to be exactly matching the firm's losses from decreasing the capacity for the other class in the *dominating interval*. Therefore, the firm's revenue is the same for any passing levels in the *dominating interval*. We also show that the firm's revenue in the *dominating interval* is equal to the revenue of the *Benchmark* case. Hence, the firm does not benefit from offering an additional test when the agent skills become perfectly correlated.

THEOREM 4. 1. *If* $\rho_B < 1/2$*, then we have that* $\lim_{\eta\to\infty}\omega_A^* = \rho_B$*,* $\lim_{\eta\to\infty}\omega_B^* = 1-\rho_B$*. Furthermore, we have that* $\lim_{\eta\to\infty}\Delta_i^* = \rho_B(1-\rho_B)$ *for all* $i \in \{A,B\}$*.*

2. *If* $\rho_B \geq 1/2$*, then we have that* $\lim_{\eta\to\infty}\omega_A^* = \lim_{\eta\to\infty}\omega_B^* = 1/2$*. Furthermore, we have that* $\lim_{\eta\to\infty}\Delta_i^* = 1/4$ *for all* $i \in \{A,B\}$*.*

3. $\lim_{\eta\to 1}\Delta_i^* = 0$ *for all* $i \in \{A,B\}$*, i.e, offering an additional test does not improve the revenue of the firm when* η *approaches 1.*

The above theorem shows the limiting behavior of the revenue improvements after offering an additional exam as the agent skills become independent. We perform a numerical study to illustrate this limiting behavior and show that more than half of the asymptotical revenue improvement stated in Theorem 4.2 can be achieved even when η is as low as 10. We set the demand rate of class A to 1 and consider three different values for the demand rate of class B in our numerical study, as presented in Figure 4. This figure also shows that benefits from offering one exam become more profound as the correlation between an agent's skills weakens, which happens as η increases.

One of the major implications of the above theorem is that the optimal passing level is one of the end points of the *dominating interval* for large values of η when the lower demand rate is less than 1/2. The firm optimally chooses the end point which allocates enough capacity to match the demand from the class with the lower demand rate. On the other hand, the optimal passing level is an interior solution when the lower demand rate is higher than 1/2. The firm divides the service capacity equally between both classes under this interior solution. We also numerically observe that exam A outperforms exam B despite the fact that they generate the same revenue in the limit. Based on these observations, we propose a heuristic solution to the firm's problem under the *One*

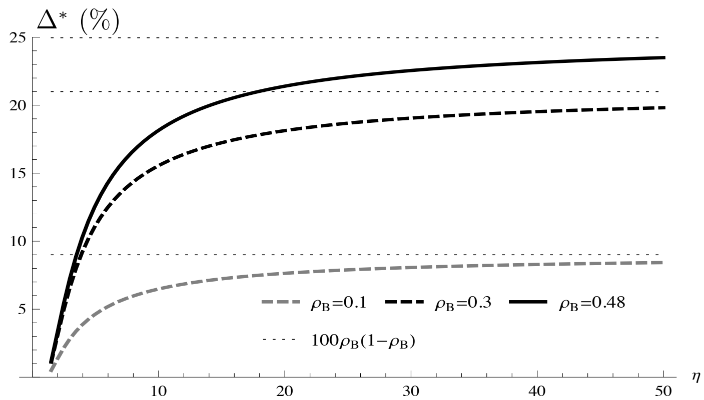


Figure 4 **The revenue improvements after offering one test as η grows.**

Test case: The firm offers exam A and sets the passing level to $F_\eta^{-1}(\rho_B)$ if $\rho_B < 1/2$ and $F_\eta^{-1}(1/2)$ otherwise. As a direct implication of Theorem 4, the heuristic solution we propose is asymptotically optimal, i.e. as $\eta \to \infty$. Furthermore, we numerically observe that our heuristic solution performs quite well. In fact, the revenue gap between the proposed heuristic and the optimal solution is below 0.2% even for low levels of the shape parameter η.

6.3. Two Skill Tests

As our final case, we now study the *Two-Tests* case. In this case, the firm sets strictly positive passing levels on both exams. Therefore, the firm's objective, in the *Two-Tests* case, is to maximize the revenue, $\Pi(\omega_A, \omega_B)$, given the constraints of $\omega_A > 0$ and $\omega_B > 0$. We denote the firm's optimal profit by Π^{**} and the optimal passing level in exam $i \in \{A, B\}$ by ω_i^{**}. We also define the relative improvement in revenue from the *One-Test* case to the *Two-Tests* case as $\Pi^{**}/\max\{\Pi_A^*, \Pi_B^*\} - 1$ and denote it by Δ^{**}.

Unlike the previous case where the skill-mix structure is only one type, the marketplace in the *Two-Tests* case can be an *N-Network*, an *M-Network*, or a *V-Network*. This will make the firm's problem more challenging than the *One-Test* case. Therefore, similar to the *One-Test* case, we will focus on the limiting cases where the agent skills are (i) perfectly correlated and (ii) independent.

When the skills of an agent are perfectly and positively correlated, we show that the firm prefers having flexible agents over dedicated agents. In a marketplace with only flexible agents, the firm has to tradeoff between the total service capacity and the equilibrium revenue of agents. It turns out, the firm strictly prefers higher capacity, and thus always improves its revenue by lowering the passing levels in both exams. This implies that the firm's revenue, after offering two exams, cannot exceed its revenues without any exams if skills are perfectly correlated. In other words, similar to the *One-Test* case, the firm does not benefit from skill screening.

In our second limiting case, we study the firm's problem when the agent skills are independent and *Uniformly* distributed. When the demand from class B (the low demand class) is greater than $1/2$, we show that none of the passing levels (ω_A, ω_B) with $\omega_A > 0$ and $\omega_B > 0$ can generate more revenue than the *One-Test* case. Hence, offering a second exam does not bring any extra benefit to the firm when $\rho_B \geq 1/2$. On the other hand, when $\rho_B < 1/2$, we identify a *dominating curve* of passing levels that generate more revenue for the firm than the rest of the passing levels do. We graphically illustrate the dominating curve in Figure 5.

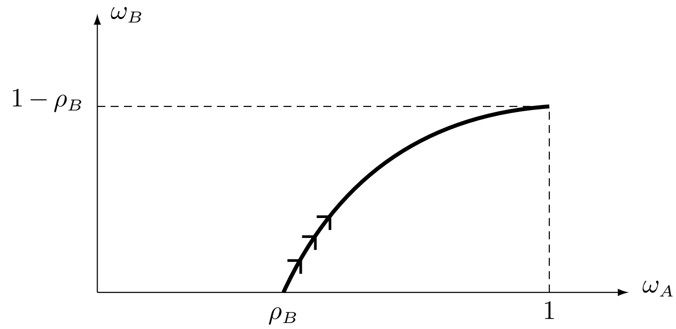


Figure 5 The dominating curve: Any (ω_A, ω_B) on the dominating curve satisfies $\omega_B = 1 - \rho_B/\omega_A$

As the firm moves on the *dominating curve* in the direction indicated by arrows (which could be achieved by increasing both passing levels), the firm faces a tradeoff: The average skill of a dedicated agent serving class B increases, which leads to higher revenues from class B, whereas the service capacity allocated to this class decreases, which means lower revenues from class A. We show that the gains from class B may outweigh the losses from class A as the firm move along the dominating curve. Thus, there may be an interior optimal solution on the *dominating curve*, and this implies that the firm is strictly better off running the second exam. However, we show that the relative improvement after the second exam cannot exceed 1.7% and is positive only when $\rho_B < 1/2$. We formally present these findings in the following theorem.

THEOREM 5. 1. *If $\rho_B < 1/2$, then we have that* $\lim_{\eta\to\infty}(\omega_A^{**}, \omega_B^{**}) = (\tilde{\omega}, 1-\rho_B/\tilde{\omega})$, *where* $\tilde{\omega} = \min\{(\rho_B^2/2)^{1/3}, (1-\sqrt{1-4\rho_B}\}$. *Furthermore, we have that* $\lim_{\eta\to\infty}\Delta^{**} \leq 1.7\%$.

2. *If $\rho_B \geq 1/2$, then we have that* $\lim_{\eta\to\infty}\Delta^{**} = 0$.

3. $\lim_{\eta\to 1}\Delta^{**} = 0$, *i.e., offering an additional test does not improve the revenue of the firm when η approaches 1.*

The above theorem only provides an upper bound for the benefits obtained from the second exam as $\eta \to \infty$. The exact expression, which is omitted for brevity, can be found in the proof of the theorem. We illustrate the relative improvement after the second exam as the skills become independent, $\lim_{\eta\to\infty} \Delta^{**}$, in Figure 6. As this figure shows, the improvements exceed 1% only for a small interval of class B demand rate, ρ_B.

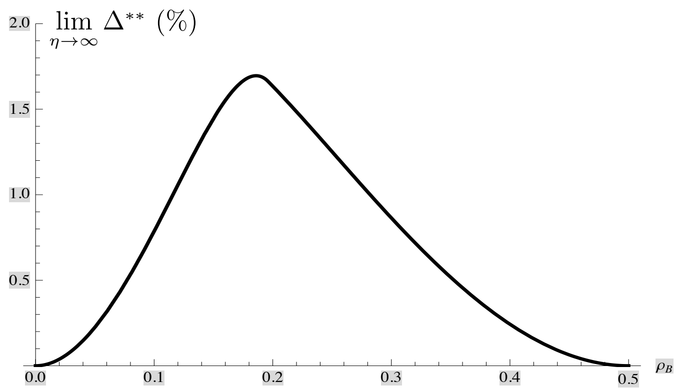


Figure 6 **Relative revenue improvements from the One-Test case to the Two-Tests case when $\eta = \infty$.**

Similar to the *One-Test* case, we perform a numerical study to illustrate the limiting behavior proven in Theorem 5. Unlike the Δ^*, Δ^{**} is not always increasing in the shape parameter η as shown in the following figure. This non-monotone structure occurs because Δ^{**} captures the revenue improvements compared to the *One-Test* case. We verified that the relative improvement in revenue from the *Benchmark* case to the *Two-Tests* case increases as the correlation between agent skills declines (i.e., as η increases).

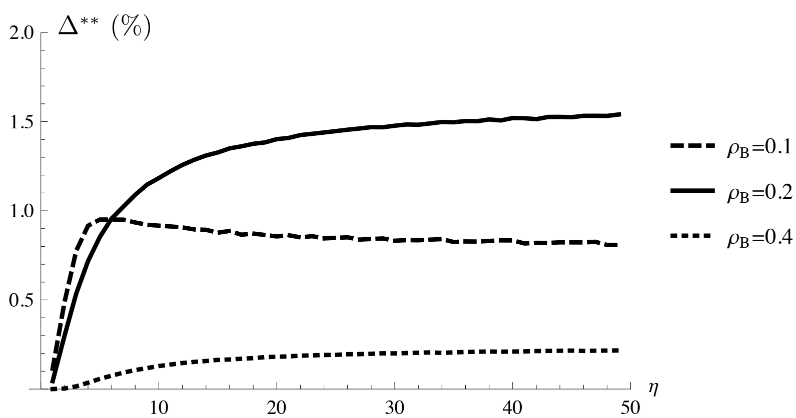


Figure 7 **Relative revenue improvements from the One-Test case to the Two-Tests case as a function of η.**

6.4. The Optimal Number of Tests

In the previous subsections, we study the firm's problem by fixing the number of skill tests it offers. We show that the revenue improvements from offering skill tests highly depend on the correlation between the skills of an agent. Specifically, we find that offering a skill test slightly improves the firm's revenue when the agent skills are highly correlated. On the other hand, if the correlation between skills is negligible, the first additional test can lead to substantial revenue improvements (as much as 25%). However, the second exam cannot generate similar levels of significant revenue improvements. We also numerically show that the firm benefits from the tests more as the skills become less correlated.

When the firm does not incur any costs for preparing and offering a skill test, the *Two-Tests* case is naturally the best option for the firm because the firm can set the passing levels to zero if needed. In the absence of skill screening costs, our findings help the moderating firm to choose the optimal level of difficulty in each exam. For instance, our results suggest that the firm should let all candidate agents pass both exams when skills are perfectly correlated, while setting strictly positive passing levels in both exams is optimal if the skill correlation can be disregarded. However, it is not unrealistic to consider a preparation and implementation cost associated with the skill tests. Then, the firm can use our findings to choose the optimal number of tests to be offered. As we show in the following corollary, letting C be the cost for offering a skill test, we can find a critical level of shape parameter $\underline{\eta}$ and show that the firm optimally offers no tests when $\eta < \underline{\eta}$. Similarly, we can show that it is optimal for the firm to offer only one test for sufficiently low levels of skill correlation as long as C is greater than a small percentage (less than 1.7%) of the total revenue.

COROLLARY 1. *For any $C > 0$, there exists a $\underline{\eta}$ such that it is optimal for the firm to offer zero tests for any $\eta < \underline{\eta}$. Furthermore, if $C/\Pi^{**} > \lim_{\eta\to\infty} \Delta^{**}$ and $C/\Pi^{o} < \lim_{\eta\to\infty} \Delta^{*}$, then there exists a $\overline{\eta}$ such that it is optimal for the firm to offer only one test for any $\eta > \overline{\eta}$.*

7. Conclusion

In this paper, we study a marketplace in which many small service providers compete with each other in providing service to two groups of self-interested customers. Service providers are distinguished with respect to their service skills, and each group of customers has different needs. An important aspect of these marketplaces that our model captures is that customers cannot learn the skills of a provider before the completion of the service. However, the moderating firm, which sets up the marketplace, may help customers by providing them with further information about the ability of candidate agents through a skill screening mechanism. Such a screening mechanism consists of skill tests determining whether or not a candidate agent is eligible to serve customers. Skill screening also helps the firm to create different skill-mix structures in the marketplace.

The main focus of this paper is to gain insights about how the moderating firm can use skill screening as a tool to maximize its revenue. Hence, we study a problem where the firm can offer two skill tests and choose passing levels in the tests it offers. As the online marketplaces we review usually attract service providers with complementary skills, we consider a family of agent skill distributions where the correlation between skills ranges from 1 (perfect correlation) to 0 (independence). We show that the level of correlation between agent skills plays a crucial role on how the firm uses the skill tests. For instance, when the agent skills are highly correlated, skill screening can hardly improve the firm's revenue. Considering possible costs associated with test preparation and execution, this result suggests that the firm is better off not offering any skill tests in marketplaces with highly compatible skills. On the other hand, we show that the firm starts to obtain considerable benefits from skill screening as the skill correlation softens. It turns out the firm may improve its revenue by as much as 25% after running skill tests in marketplaces where the skill correlation is low. We also show that the firm does not need to run both of the exams to achieve these high levels of benefits from testing, even when the screening costs are small (e.g. 1.7% of the firm's revenue).

As we mentioned before, one can view skill screening as a tool to regulate the marketplace. Thus, our results also shed light on the relationship between the level of skill correlation and how much intervention the marketplace requires. In particular, we show that the higher the correlation between the skills of an agent is, the less regulation/intervention the firm needs. Furthermore, we recommend that moderating firms not use both of the available exams when regulation is necessary.

In this paper, we study a marketplace where the service capacity is scarce and customers' outside utility is normalized to zero. Our key findings, which explain the relationship between skill screening and correlation, would continue to hold when the capacity is ample and customers have a positive outside utility. In the case of ample capacity (i.e., when $\rho_A + \rho_B < 1$), the firm has to choose strictly positive passing levels because otherwise, service providers would charge very low prices due to intensified competition. Similarly, when the customers have a strictly positive outside utility, it is profitable for the firm to increase the passing levels to match this outside option. In fact, relaxing these two assumptions will shift the origin of the feasible passing level space outwards from $(0,0)$ to $(\underline{\omega}_A, \underline{\omega}_B)$, where both passing levels are positive. Thus, the firm always offers two tests, even when they are costly. However, the remaining important question is how much benefit the firm can obtain when it further intervenes in the marketplace by making the skill tests more comprehensive than the shifted origin. After relaxing the aforementioned assumptions, we show that the firm does not set passing levels higher than the shifted origin when skills are perfectly correlated. In other words, more intervention is not profitable for the firm under perfectly correlated skills. Furthermore, the firm obtains significant benefits from making one of the exams more comprehensive compared to the new origin $(\underline{\omega}_A, \underline{\omega}_B)$ when the correlation between skills loosens. Similar to our results in Section 6, this implies that intervening in the marketplace via only one exam can lead to sizable revenue gains for the firm. However, setting the passing levels higher than $(\underline{\omega}_A, \underline{\omega}_B)$ in both exams only leads to small gains, which again establishes that running the second exam is not advisable for the moderating firm when skill screening is costly.

References

Allon, G., A. Bassamboo, E. B. Çil. 2012. Large-scale service marketplaces: The role of the moderating firm. *Management Sci.* **58**(10) 1854–1872.

Allon, G., A. Federgruen. 2007. Competition in service industries. *Operations Research* **55**(1) 37–55.

Allon, G., I. Gurvich. 2010. Pricing and dimensioning competing large-scale service providers. *Manufacturing & Service Operations Management* **12**(3) 449–469.

Amihud, Y., H. Mendelson. 1980. Dealership market: Market-making with inventory. *Journal of Financial Economics* **8**(1) 31–53.

Bassamboo, A., R. S. Randhawa, J. A. Van Mieghem. 2010. Optimal flexibility configurations in newsvendor networks: Going beyond chaining and pairing. *Management Sci.* **56**(8) 1285–1303.

Biais, B., L. Glosten, C. Spattt. 2005. Market microstructure: A survey of microfoundations, empirical results, and policy implications. *Journal of Financial Markets* **8**(2) 217–264.

Bish, E. K., Q. Wang. 2004. Optimal investment strategies for flexible resources, considering pricing and correlated demands. *Operations Research* **52** 954–964.

Burdett, K., D. Mortensen. 1998. Wage differentials, employer size, and unemployment. *International Economic Review* **39**.

Cachon, G., P. T. Harker. 2002. Competition and outsourcing with scale economies. *Management Sci.* **48**(10) 1314–1333.

Callen, J. L., B. Sarath. 1995. Risk aversion, generalized correlation and investment in manufacturing capacities. *Journal of Operations Management* **12** 225–238.

Chen, H., Y. Wan. 2003. Price competition of make-to-order firms. *IIE Transactions* **35**(9) 817–832.

Chen, Y., C. Maglaras, G. Vulcano. 2008. Design of an aggregated marketplace under congestion effects: Asymptotic analysis and equilibrium characterizatio. Working Paper.

Chod, J., N. Rudi. 2005. Resource flexibility with responsive pricing. *Operations Research* **53** 532–548.

Fine, C. H., R. M. Freund. 1990. Optimal investment in product-flexible manufacturing capacity. *Management Science* **36** 449–466.

Garman, M. 1976. Market microstructure. *Journal of Financial Economics* **3**(3) 257–275.

Goyal, M., S. Netessine. 2007. Strategic technology choice and capacity investment under demand uncertainty. *Management Science* **53**(2) 192–207.

Gupta, D., Y. Gerchak, J. A. Buzacott. 1992. The optimal mix of flexible and dedicated manufacturing capacities: Hedging against demand uncertainty. *International Journal of Production Economics* **28** 309–320.

Ho, T., H. Stoll. 1983. The dynamics of dealer markets under competition. *Journal of Finance* **38**(4) 1053–1074.

Manning, A. 2004. Monopsony and the efficiency of labour market interventions. *Labour Economics* **11**(2) 145–163.

Michaelides, M. 2010. Labour market oligopsonistic competition: The effect of worker immobility on wages. *Labour Economics* **17**(1).

Moreno, D., J. Wooders. 1996. Coalition-proof equilibrium. *Games and Economic Behavior* **17** 80–112.

Naor, P. 1969. The regulation of queue size by levying tolls. *Econometrica* **37**(1) 15–24.

Rath, K. P. 1992. A direct proof of the existence of pure strategy equilibria in games with a continuum of players. *Economic Theory* **2**(3) 427–433.

Ray, I. 1996. Coalition-proof correlated equilibirum: A definition. *Games and Economic Behavior* **17** 56–79.

Van Mieghem, J. A. 1998. Investment strategies for flexible resources. *Management Science* **44** 1071–1078.

APPENDICES

We provide the proofs of the supplementary results used in our proofs in the Supporting Document. The Supporting Document also presents the extended proofs of our results.

Appendix A: Proofs in Section 5

A.1. Supplementary result for the proof of Proposition 1

LEMMA 1. *Given any* $(\mathbf{r_A}, \mathbf{y_A})$*, let* $D^{ce}_{A_n}$ *be the fraction of customers choosing sub-pool* A_n *in the customer equilibrium and* U^{ce}_A *be the corresponding expected utility of customers from the sub-pools attracting positive demand.*

1. *If* $U^{ce}_A = r_{A_\ell}$ *for some* $\ell \in \{1, \dots, N_A\}$*, then we have that* $\rho^0_\ell \leq \rho_A \sum_{n=1}^{N_A} D^{ce}_{A_n} \leq \rho^0_\ell + y_{A_\ell}$.
2. *If* $U^{ce}_A < r_{A_\ell}$ *for some* $\ell \in \{1, \dots, N_A\}$*, then we have that* $\rho_A \sum_{n=1}^{N_A} D^{ce}_{A_n} > \rho^0_\ell + y_{A_\ell}$.
3. *If* $r_{A_\ell} < U^{ce}_A < r_{A_{\ell-1}}$ *for some* $\ell \in \{2, \dots, N_A\}$*, then we have that* $\rho_A \sum_{n=1}^{N_A} D^{ce}_{A_n} < \rho^0_\ell$.
4. *If* $U^{ce}_A > r_{A_\ell}$ *for some* $\ell \in \{2, \dots, N_A\}$*, then we have that* $\rho_A \sum_{n=1}^{N_A} D^{ce}_{A_n} < \rho^0_\ell$.

Here, we let $\rho^0_\ell = \left(\sum_{n=1}^{\ell-1} r_{A_n} y_{A_n}\right)/r_{A_\ell}$*, and* $\rho^0_1 = 0$.

The proof of the above lemma can be found in Appendix S.1.

A.2. Proof of Proposition 1

Below, we state and prove an extended version of Proposition 1. Proposition 1 is a direct implication of Proposition 3. When $\rho_i > \rho_\ell^0 + y_{i_\ell}$, the result holds since $\rho D_{i_\ell}^{ce} > y_{i_\ell}$ by parts 2 and 3. When, $\rho_i \in [\rho_\ell^0, \rho_\ell^0 + y_{i_\ell}]$, this time the result holds by part 1. Finally, when $\rho_i < \rho_\ell^0$, the result holds again by part 1.

Proposition 3. *Given any $(\mathbf{r_A}, \mathbf{y_A})$, let $D_{A_n}^{ce}$ be the fraction of customers choosing sub-pool A_n in a customer equilibrium, we obtain the following:*

1. *If $\rho_A \in [\rho_\ell^0, \rho_\ell^0 + y_{A_\ell}]$ for some $\ell \in \{1, \dots, N_A\}$, then we have that $D_{A_n}^{ce} = r_{A_n} y_{A_n} / r_{A_\ell} \rho_A$ for any $n < \ell$, and $\sum_{n \geq \ell} D_{A_n}^{ce} = (\rho_A - \rho_\ell^0)/\rho_A$.*
2. *If $\rho_A \in (\rho_\ell^0 + y_{A_\ell}, \rho_{\ell+1}^0)$ for some $\ell \in \{1, \dots, N_A - 1\}$, then we have that $D_{A_n}^{ce} = r_{A_n} y_{A_n} / \left(\sum_{m=1}^{\ell} r_{A_m} y_{A_m}\right)$ for any $n \leq \ell$, and $D_{A_n}^{ce} = 0$ for any $n > \ell$.*
3. *If $\rho_A \in (\rho_{N_A}^0 + y_{A_{N_A}}, \infty)$, then we have that $D_{A_n}^{ce} = r_{A_n} y_{A_n} / \max\left\{\sum_{m=1}^{N_A} r_{A_m} y_{A_m}, \rho_A \underline{u}\right\}$ for any $n \leq N_A$.*

Proof: We first show that the expected utility of customers in the equilibrium, say U_A^{ce}, is r_{A_ℓ} in part 1. To do this, we first suppose $U_A^{ce} < r_{A_\ell}$. Then, by Lemma 1.2, we have that $\sum_{n=1}^{N_A} D_{A_n}^{ce} > 1$. However, this is clearly a contradiction since $\sum_{n=1}^{N_A} D_{A_n}^{ce}$ cannot exceed 1. Similarly, $U_A^{ce} > r_{A_\ell}$ leads to a contradiction because it implies that $\sum_{n=1}^{N_A} D_{A_n}^{ce} < 1$ despite customer utility is strictly positive. As a direct implication of $U_A^{ce} = r_{A_\ell}$, we have $\rho_A D_{A_1}^{ce} = r_{A_1} y_{A_1} / r_{A_\ell}$. Furthermore, sub-pool-1 serves only its customers. Therefore, customers picking sub-pool-2 can only get their service from sub-pool-2, and because of that we should $\rho_A D_{A_2}^{ce} = r_{A_2} y_{A_1} / r_{A_\ell}$. We can apply this argument to any sub-pool n for $n < \ell$, and conclude that $\rho_A D_{A_n}^{ce} = r_{A_n} y_{A_n} / r_{A_\ell}$ and sub-pool n serves only its own customers for any $n < \ell$. Finally, since $\rho_A - \rho_A \sum_{n=1}^{\ell-1} D_{A_n}^{ce} = \rho_A - \rho_\ell^0 \leq y_{A_\ell}$, we should have that $\sum_{n=1}^{N_A} D_{A_n}^{ce} = 1$. Thus, we have that $\sum_{n=\ell}^{N_A} D_{A_n}^{ce} = (\rho_A - \rho_{A_\ell}^0)/\rho$. The proofs of parts 2 and 3 is very similar to Part 2 and can be seen in Appendix S.2 in the Supporting Document.

A.3. Proof of Theorem 1

In this proof, we, first, focus on symmetric *Market Equilibrium*, where the same type of agents (dedicated or flexible) charge the same price. We let $(p_D, p_F; \alpha_D, \alpha_F)$ be a symmetric *Market Equilibrium* under the special market structure. We will provide a discussion on non- symmetric *Market Equilibrium* at the end of the proof. As a first step towards characterizing the symmetric equilibrium, we derive the revenue of agents when all the agents in the same pool charge the same price as follows:

Corollary 2. *Let $V_D^{ce}(p_D, p_F)$ and $V_F^{ce}(p_D, p_F)$ be the revenue of a dedicated and a flexible agent, respectively, when all the dedicated agents charge p_D and all flexible agents charge p_F. If $R_D - p_D \neq R_F - p_F$, then we have that*

$$V_h^{ce}(p_D, p_F) = \begin{cases} p_h (\rho/\alpha_h)\tau & \text{if } \rho \leq \alpha_h \\ p_h \tau & \text{if } \rho > \alpha_h \end{cases}, \text{ and } V_l^{ce}(p_D, p_F) = \begin{cases} 0 & \rho < \rho_l^0 \\ p_l(\rho - \rho_l^0)/\alpha_l \tau & \rho \in [\rho_l^0, \rho_l^0 + \alpha_l] \\ p_l \tau & \rho > \rho_l^0 + \alpha_l, \end{cases}$$

where $\rho_l^0 = \frac{R_h - p_h}{R_l - p_l}\alpha_h$, $h = \arg\max_{i \in \{D,F\}}(R_i - p_i)$, and $l = \arg\min_{i \in \{D,F\}}(R_i - p_i)$.

Since $(p_D, p_F; \alpha_D, \alpha_F)$ is an equilibrium, there exists a sequence of (ϵ^k, δ^k)-*ME*, say $(p_D^k, p_F^k; \alpha_D, \alpha_F)$, such that $(p_D, p_F) = \lim_{k\to\infty}(p_D^k, p_F^k)$ where $\lim_{k\to\infty}(\epsilon^k, \delta^k) = (0,0)$. We let $V_D^{sm}(k)$ and $V_F^{sm}(k)$ be the revenue of a dedicated and a flexible agent, respectively, according to $(p_D^k, p_F^k; \alpha_D, \alpha_F)$. Then, we have that $V_i^{sm} = \lim_{k\to\infty} V_i^{sm}(k)$ for all $i \in \{D, F\}$. We suppose $R_D \geq R_L$ without loss of generality.

1 We show that $V_D^{sm} = V_F^{sm} = 0$ by contradiction. Thus, we suppose that either $V_D^{sm} > 0$ or $V_F^{sm} > 0$ is true on the contrary and find a contradiction for any possible price pair (p_D, p_F) satisfying either of these conditions. To this end, we follow a case-by-case analysis:

i. $(\mathbf{R_D - p_D = R_F - p_F})$**:** Notice to that dedicated agents are under-utilized since $\rho < \alpha_D$, and thus their revenue is at most $\tau(\rho p_D^k)/\alpha_D$. When a small fraction of dedicated agents deviate and cut their prices by an arbitrarily small ζ, the revenue of deviating agents will be $\tau(p_D^k - \zeta)$ for large k by Proposition 1. As this deviation improves the revenue of deviating agents, any (p_D, p_F) satisfying $R_D - p_D = R_F - p_F$ cannot emerge as an equilibrium price pair.

ii. $(\mathbf{R_F - p_F > R_D - p_D})$**:** Let $\rho_D^0 = \frac{R_F - p_F}{R_D - p_D}\alpha_F$. In this case, we have two sub-cases: **a)** $(\rho \leq \boldsymbol{\rho_D^0})$**:** By Corollary 2, we have that $\lim_{k\to\infty} V_D^{sm}(k) = 0$ since $\rho \leq \rho_D^0$. We also should have that $V_F^{sm} > 0$, which implies that $p_F > 0$. Then, when a small fraction of dedicated agents deviate and charge p' with $0 < p' < p_F$, the revenue of deviating agents will be p' for large k by Proposition 1. **b)** $(\rho > \boldsymbol{\rho_D^0})$**:** By Corollary 2, we have that $\lim_{k\to\infty} V_D^{sm}(k) < \tau p$. Then, as in Part 1.i, a small group of dedicated agents can improve their revenues by cutting their price.

iii. $(\mathbf{R_D - p_D > R_F - p_F})$**:** By Corollary 2, we have that $\lim_{k\to\infty} V_D^{sm}(k) = \rho p_D$ and $\lim_{k\to\infty} V_F^{sm}(k) = 0$. We should have that $p_D > 0$ to make sure $V_D^{sm} > 0$. Then, as in **Part 1.i**, a small group of dedicated agents can improve their revenues by cutting their price.

2. We suppose that either $V_D^{sm} > \tau(R_D - R_F)$ or $V_F^{sm} > 0$ on the contrary and follow a case-by-case analysis:

i. $(\mathbf{R_D - p_D = R_F - p_F})$**:** Notice that the utilization of at least one group of agents, say group-i where $i \in \{D, F\}$, should be less than the demand rate over total capacity, i.e. $\rho/(\alpha_D + \alpha_F)$. Then, when a small fraction of group-i agents deviate and cut their prices by an arbitrarily small ζ, the revenue of deviating agents will be $p_i^k - \zeta$ for large k by Proposition 1.

ii. $(\mathbf{R_D - p_D > R_F - p_F})$**:** By Corollary 2, we have that $\lim_{k\to\infty} V_D^{sm}(k) = \tau p_D$ and $\lim_{k\to\infty} V_F^{sm}(k) = 0$. We should also have that $p_D > R_D - R_F$ to make sure that $V_D^{sm} > \tau(R_D - R_F)$. Then, when a small fraction of flexible agents deviate and charge p' with $0 < p' < (R_F - R_D + p_D)$, the revenue of deviating agents will be p' for large k by Proposition 1.

The proofs for parts 3, 4 and 5 are similar to the previous ones.

Non-symmetric equilibria: We do not rule out the existence of a non-symmetric equilibrium outcome, where the same type of agents (dedicated or flexible) charge different prices. However, we show that the possibility of non-symmetric equilibrium can be ignored using the results of Proposition 4 as we focus on agent revenues. The following proposition proves that the same type of agents serving the same customer class in any non-symmetric *Market Equilibrium* must earn zero revenue.

Proposition 4. *Given any Market Equilibrium* $(\mathbf{r_i}, \mathbf{y_i})$*, let* $V^{me}_{i_nD}$ *and* $V^{me}_{i_nF}$ *be the equilibrium revenue of a dedicated and a flexible agents in sub-pool* i_n *for all* $i \in \{A, B\}$*, respectively.*

1. *If the number of different prices announced by the dedicated (flexible) agents serving class* $i \in \{A, B\}$ *is two or more, then we have that* $V^{me}_{i_nD} = 0$ $(V^{me}_{i_nF} = 0)$ *for all* $n \in \{1, \ldots, N_i\}$*.*
2. $V^{me}_{F_{A_n}} = V^{me}_{F_{B_n}}$ *for any* $n \in \{1, \ldots, N_A\}$*,* $m \in \{1, \ldots, N_B\}$*.*

The above proposition directly implies that any non-symmetric equilibrium does not affect our results for the revenue of dedicated agents in parts 1 and 2 and for the revenue of flexible agents in parts 1-4 because we do not exclude the possibility of zero revenue in these cases. In the remaining cases, we can show that there is not any non-symmetric equilibrium as follows: Suppose there is a non-symmetric equilibrium, where dedicated agents charge different prices, when $\rho > \alpha_D$ and $R_D > R_F$. By Proposition 4, we should have that all of the dedicated agents earn zero in the equilibrium. However, a small group of dedicated agents can guarantee a strictly positive revenue by charging a very low price ζ since $\rho > \alpha_D$. Similarly, we can rule out any non-symmetric equilibrium where flexible agents charge different prices if $\rho > 1$.

Proposition 4 also shows that the equilibrium revenues of the flexible agents must be the same regardless of the customer class they serve. Hence, Proposition 4 establishes that the same type of agents earn the same revenue in any *Market Equilibrium*. The proof of Proposition 4 can be seen in in Appendix S.3 in the Supporting Document.

We provide an extended proof of this theorem and a discussion about the existence of the *Market Equilibrium* in Appendix S.4.

Appendix B: Proofs in Section 6

Note that the service capacity τ *is only a multiplier for the revenues of the agents. Therefore, we normalize the service capacity to 1 for notational convenience in our proofs.*

B.1. Proof of Proposition 2

Using the standard definition of the correlation coefficient, for any given shape parameter η, we have that

$$Corr(S_A, S_B) = \frac{\eta(\eta+3)(3\eta+1)(\eta(4\eta+9)+4)}{4(\eta+1)^2(\eta(\eta+1)(\eta(\eta+3)+4)+1)}$$

As $\eta \to \infty$, we have that $Corr(S_A, S_B) \to 0$ because the denominator is a higher degree polynomial than the numerator is. On the other hand, as $\eta \to 1$, we have that $Corr(S_A, S_B) \to \frac{4\times4\times(13+4)}{4\times4\times(2\times8+1)} = 1$.

B.2. Proof of Theorem 2

Let γ^o_A be the fraction of flexible agents serving class A. Then, we must have that $1 - \rho_B \le \gamma^o_A \le \rho_A$ because otherwise flexible agents serving one of the classes would earn zero according to Theorem 1. Furthermore, Theorem 1 implies that all flexible agents earn $\mathbf{E}[S_\eta]$. Finally, our claim holds because

$$\mathbf{E}[S_\eta] = \frac{\eta+1}{\eta-1}\int_0^1 \int_{s_A^\eta}^{s_A^{1/\eta}} s_B ds_A ds_B = \frac{\eta+1}{2(\eta-1)}\int_0^1 (s_A^{2/\eta} - s_A^{2\eta}) ds_A = \frac{(\eta+1)^2}{(\eta+2)(2\eta+1)}.$$

B.3. Proof of Theorem 3

We suppose the firm offers exam A throughout the proof. We first show the $h(\omega) \equiv \omega^{1/\eta} + \omega^{\eta} - 2\omega = 0$ has a unique non-trivial solution $\bar{\omega} \in (0,1)$. We have that $h''(\omega) < 0$ for any $\omega < \eta^{-\frac{3\eta}{\eta^2-1}} \in (0,1)$ and $h''(\omega) \geq 0$ otherwise. Combining this with the facts that $h'(0) > 0$ and $h'(1) > 0$, we can find two critical levels of ω, ω_1 and ω_2 with $0 < \omega_1 < \omega_2 < 1$, such that $h'(\omega) < 0$ for any $\omega \in (\omega_1, \omega_2)$ and $h'(\omega) \geq 0$ otherwise. Then, using $h(0) = h(1) = 0$, we have that $h(\omega) > 0$ for any $\omega \leq \omega_1$, $h(\omega) < 0$ for any $\omega \geq \omega_2$. Furthermore, $h(\omega)$ is decreasing in ω for any $\omega \in (\omega_1, \omega_2)$. Thus there exists a unique $\bar{\omega} \in (\omega_1, \omega_2)$ with $h(\bar{\omega}) = 0$ and $h(\omega) > 0$ for any $\omega < \bar{\omega}$ and $h(\omega) < 0$ for any $\omega > \bar{\omega}$.

We next prove that the interval $[F_\eta^{-1}(1-\rho_A), F_\eta^{-1}(\rho_B)]$ is the *dominating interval*. For all passing levels with $\omega > F_\eta^{-1}(\rho_B)$, dedicated agents earn zero equilibrium revenue according to Theorem 1 because $\alpha_B > \rho_B$. Furthermore, all flexible agents serve class A and charge R_{AF}. Therefore, the total revenue of the marketplace is $\Pi(\omega, 0) = \int_\omega^1 \int_{s_A^\eta}^{s_A^{1/\eta}} s_A f_{A,B} ds_A ds_B$. Notice that $\Pi(\omega,0)$ is decreasing in ω, so that $\Pi(\omega,0) < \Pi(F_\eta^{-1}(\rho_B), 0)$ for any $\omega > F_\eta^{-1}(\rho_B)$.

Now, consider the passing levels with $\omega < F_\eta^{-1}(1-\rho_A)$. The total capacity of flexible agents is above both ρ_A and ρ_B when $\omega < F_\eta^{-1}(1-\rho_A)$. Therefore, their revenue in the equilibrium cannot exceed $\min\{R_{AF}, R_{BF}\}$ because flexible agents must earn the same equilibrium revenue regardless of the class they serve by Proposition 4.2. We show that $R_{AF} \geq R_{BF}$: $R_{AF} = R_{BF}$ at $\omega = 0$ and $\omega = 1$. Moreover, $(R_{AF} - R_{BF})\alpha_F$ is increasing in ω for any $\omega < \bar{\omega}$ and decreasing otherwise its derivative with respect to ω is $\frac{(\eta+1)(\omega^{1/\eta} - \omega^{\eta})}{2(\eta-1)} h(\omega)$. Therefore, all flexible agents earn at most R_{BF}. We also have that dedicated agents earn at most R_B. Combining these two, we have that $\Pi(\omega, 0) \leq \mathbf{E}[S_\eta] \leq \Pi(F_\eta^{-1}(1-\rho_A), 0)$ where the last inequality holds because $R_{AF} \geq R_{BF}$ and all flexible agents can serve class A when $\omega = F_\eta^{-1}(1-\rho_A)$.

After proving $[F_\eta^{-1}(1-\rho_A), F_\eta^{-1}(\rho_B)]$ dominates the rest of the passing levels, we focus on the total revenue in the *dominating interval*. For any $\omega \in [F_\eta^{-1}(1-\rho_A), F_\eta^{-1}(\rho_B)]$, flexible agents sustain an equilibrium by serving only class A and charging R_{AF} because $\alpha_F \leq \alpha_A$. Moreover, the dedicated agents can charge R_B since $\alpha_B \leq \rho_B$. Therefore, the total revenue or any $\omega \in [F_\eta^{-1}(1-\rho_A), F_\eta^{-1}(\rho_B)]$ is $\Pi(\omega, 0) = \int_0^\omega \int_{s_A^\eta}^{s_A^{1/\eta}} s_B f_{A,B} ds_A ds_B + \int_\omega^1 \int_{s_A^\eta}^{s_A^{1/\eta}} s_A f_{A,B} ds_A ds_B$. Taking the derivative of the revenue function, we have that $\Pi'(\omega, 0) = \frac{(\eta+1)(\omega^{1/\eta} - \omega^{\eta})}{2(\eta-1)} h(\omega)$. Therefore, the optimal passing level is the solution of $h(\omega) = 0$ if $\bar{\omega}$ is inside the dominating interval. Otherwise, one of the end points of the dominating interval is optimal.

B.4. Proof of Theorem 4

1. & 2. We first want to note that $\lim_{\eta\to\infty} \bar{\rho} = 1/2$ because $\lim_{\eta\to\infty} \bar{\omega} = 1/2$ and $\lim_{\eta\to\infty} F_\eta(\omega) = \omega$. Therefore, we have that $1 - \rho_A < \rho_B < \lim_{\eta\to\infty} \bar{\rho}$. This implies that for sufficiently large η, the optimal passing level in exam A is $\min\{F_\eta(\rho_B), \bar{\omega}\}$, and thus we have that $\lim_{\eta\to\infty} \omega_A^* = \lim_{\eta\to\infty} \min\{F_\eta(\rho_B), \bar{\omega}\} = \min\{\rho_B, 1/2\}$. Similarly, $\rho_A > 1 - \rho_B > \lim_{\eta\to\infty} \bar{\rho}$ implies that $\lim_{\eta\to\infty} \omega_B^* = \max\{1 - \rho_B, 1/2\}$.

Using these limiting results, we have that $\lim_{\eta\to\infty} \Pi_A^* = \lim_{\eta\to\infty} \Pi_B^* = \int_0^{\rho_B} 1/2 ds_A + \int_{\rho_B}^1 s_A ds_A = (1 + \rho_B - \rho_B^2)k/2$ when $\rho_B < 1/2$. Then using the fact that $\lim_{\eta\to\infty} \Pi^o = k/2$, our claim about the revenue improvement holds true for any $\rho_B < 1/2$ because $\lim_{\eta\to\infty} \Delta^* = \lim_{\eta\to\infty} \Pi_A^*/\Pi^o - 1 = \rho_B(1-\rho_B)$. Similarly, for any $\rho_B \geq 1/2$, we have that $\lim_{\eta\to\infty} \Pi_A^* = \lim_{\eta\to\infty} \Pi_B^* = 5k/8$, and thus $\lim_{\eta\to\infty} \Delta^* = 1/4$.

3. In this limiting case, we have that $\lim_{\eta\to 1}\Pi(\omega,0)=\int_0^1 4s_A^2 ln(1/s_A)ds_A = 4k/9$ for any ω in the *dominating interval*. Then, our claim about the revenue improvement holds true because $\lim_{\eta\to\infty}\Pi^o = 4k/9$.

B.5. Proof of Theorem 5

1. & 2. Let $\hat{\pi}(\omega)=\omega\big(\int_{\max\{1-\rho_B/\omega,0\}}^1 sds\big)+\int_\omega^1 sds$. $\hat{\pi}(\omega)$ is concave in ω and increasing in ω for any $\omega\le\tilde{\omega}$ if $\rho_B<1/2$ where $\tilde{\omega}$ is as described in the theorem. Furthermore, $\hat{\pi}(\omega)$ is maximized at $\min\{\tilde{\omega},1/2\}$ if $\rho_B\ge 1/4$.

Using these properties of $\hat{\pi}(\omega)$, we show that $\lim_{\eta\to\infty}\Pi(\omega_A,\omega_B)$ is bounded above as the following lemma proves:

Proposition 5. $\lim_{\eta\to\infty}\Pi(\omega_A,\omega_B)\le\hat{\pi}(\tilde{\omega})$ *if* $\rho_B<1/2$ *and* $\lim_{\eta\to\infty}\Pi(\omega_A,\omega_B)\le\hat{\pi}(1/2)$ *if* $\rho_B\ge 1/2$.

As a direct implication of Lemma 5, we have that $\lim_{\eta\to\infty}\Delta^{**}=0$ when $\rho\ge 1/2$ because $\lim_{\eta\to\infty}\Pi^*=\hat{\pi}(1/2)$ when $\rho_B\ge 1/2$. Furthermore, when $\rho<1/2$, Proposition 5 implies that

$$\lim_{\eta\to\infty}\Delta^{**}=\begin{cases}\frac{\rho_B(2-3\sqrt[3]{2\rho_B}+2\rho_B)}{2+2(1-\rho_B)\rho_B} & \textit{if } \rho_B>3\sqrt{3}-5,\\ \frac{(3-\sqrt{1-4\rho_B})(2\rho_B+\sqrt{1-4\rho_B}+1)}{4+4(1-\rho_B)\rho_B}-1 & \textit{if } \rho_B\le 3\sqrt{3}-5.\end{cases}$$

We also show that $\lim_{\eta\to\infty}\Delta^{**}$ is decreasing for any $\rho_B>3\sqrt{3}-5$. Thus, $\lim_{\eta\to\infty}\Delta^{**}\le\max_{\rho_B\le 3\sqrt{3}-5}\lim_{\eta\to\infty}\Delta^{**}=0.0169563\le 1.7\%$.

Finally, the limiting behavior of the optimal passing levels $(\omega_A^{**},\omega_B^{**})$ holds true because under these passing levels all flexible agents serve class A as $\eta\to\infty$, and thus we have that $\lim_{\eta\to\infty}\Pi(\omega_A^{**},\omega_B^{**})=\hat{\pi}(\tilde{\omega})$, which proves the optimality since the upper bound is achieved.

Proof of Proposition 5: We will follow a case-by-case analysis based on the regions described in Figure 8 to prove our claim. For notational convenience, we use the upper-script ˜ to denote the limit of the revenue function, expected reward functions, and fraction of agents as $\eta\to\infty$. We also note that the customers expect the same reward from flexible and dedicated agents at the the limit. Therefore, we denote the class $i\in\{A,B\}$ customers' expected reward at the limit by $\tilde{R}_i$.

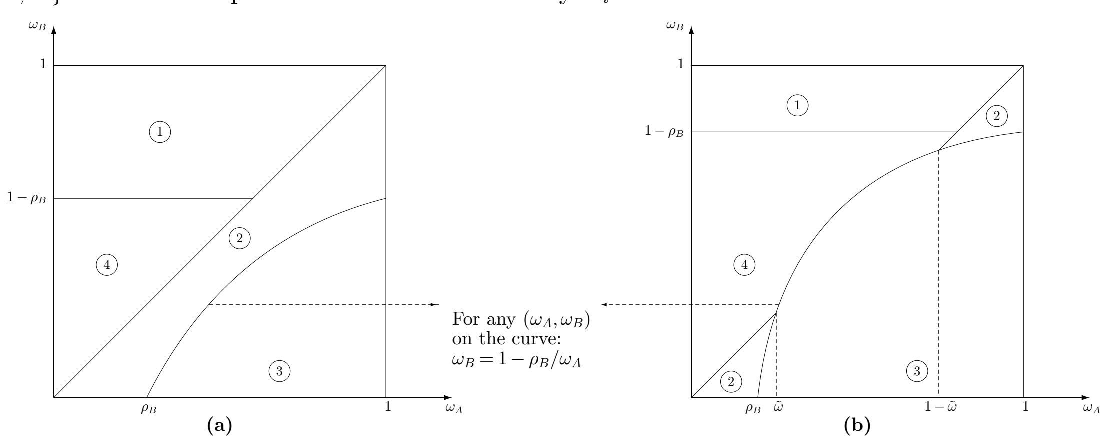


Figure 8 **Different regions that a given passing levels (ω_A,ω_B) falls as $\eta\to\infty$: (a) $\rho_B\ge 1/4$, (b) $\rho_B<1/4$.**

Region-1: For any (ω_A,ω_B) in this region, we have that $\tilde{R}_A\le\tilde{R}_B$ since $\omega_A\le\omega_B$. Therefore, the equilibrium revenue of flexible agents cannot exceed $\tilde{R}_B$. Furthermore, the equilibrium revenue of dedicated agents serving class $i\in\{A,B\}$ cannot exceed $\tilde{R}_i$. Thus, for any (ω_A,ω_B) in this region, we have that $\tilde{\Pi}(\omega_A,\omega_B)\le$

$\tilde{R}_B[\tilde{\alpha}_B + \tilde{\alpha}_F] + \tilde{R}_A\tilde{\alpha}_A = \int_{\omega_B}^1 sds + \omega_B \int_{\omega_A}^1 sds \leq \int_{\omega_B}^1 sds + \omega_B \int_0^1 sds = (1 + \omega_B - \omega_B^2)/2 = \hat{\pi}(1 - \omega_B)$. This implies that $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(\rho_B)$ when $\rho_B < 1/2$, because $\hat{\pi}(1 - \omega_B)$ is decreasing in ω_B for any $\omega_B \geq 1 - \rho_B$. Then, our claim of $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(\tilde{\omega})$ holds true because $\hat{\pi}(\omega)$ is increasing for any $\omega \leq \tilde{\omega}$ and $\rho_B \leq \tilde{\omega}$ when $\rho_B < 1/2$. On the other hand, when $\rho_B \geq 1/2$, $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(1-\omega_B)$ directly implies that $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(1/2)$ for any for any (ω_A, ω_B) in this region

Region-2: For any (ω_A, ω_B) in this region, we have that $\tilde{R}_A \geq \tilde{R}_B$ since $\omega_A \leq \omega_B$. Therefore, the equilibrium revenue of flexible agents cannot exceed $\tilde{R}_A$. Then, similar to part 1, we have that $\tilde{\Pi}(\omega_A, \omega_B) \leq \tilde{R}_A[\tilde{\alpha}_A + \tilde{\alpha}_F] + \tilde{R}_B\tilde{\alpha}_B = \int_{\omega_A}^1 sds + \omega_A \int_{\omega_B}^1 sds \leq \int_{\omega_A}^1 sds + \omega_B \int_{\max\{1-\rho_B/\omega_A\}}^1 sds = \hat{\pi}(\omega_A)$. We consider three cases to prove our claim about the upper bound of $\tilde{\Pi}(\omega_A, \omega_B)$: (i) $\rho_B \geq 1/2$, (ii) $1/4 \leq \rho_B < 1/2$, and (iii) $\rho_B < 1/4$.

When $\rho_B \geq 1/2$, $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(\omega_A)$ directly implies that $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(1/2)$ because $\omega = 1/2$ is the maximizer of $\hat{\pi}(\omega)$. Similarly, when $1/4 \leq \rho_B < 1/2$, $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(\omega_A)$ directly implies that $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(\tilde{\omega})$ because $\omega = \tilde{\omega}$ is the maximizer.

When $\rho_B < 1/4$, any (ω_A, ω_B) in this region must satisfy either $\omega_A \leq \tilde{\omega}$ or $\omega_A \geq 1 - \tilde{\omega}$ because the $\omega_B = 1 - \rho_B/\omega$ curve crosses the 45^o line as shown in Figure 8.b. Then, for any $\omega_A \leq \tilde{\omega}$, $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(\omega_A)$ implies that $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(\tilde{\omega})$ because $\hat{\pi}(\omega)$ is increasing in ω when $\omega \leq \tilde{\omega}$. We also want to note that $\hat{\pi}(\omega)$ is decreasing in ω when $\omega \geq 1 - \tilde{\omega}$ because $\hat{\pi}(\tilde{\omega})$ is maximized at $\sqrt[3]{\rho_B^2/2}$ and $\sqrt[3]{\rho_B^2/2} < 1/2 < 1 - \tilde{\omega}$. Therefore, for any $\omega_A \geq 1 - \tilde{\omega}$, $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(\omega_A)$ implies that $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(1 - \tilde{\omega})$. Combining these two upper bounds, we have that $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(\tilde{\omega})$ because $\hat{\pi}(\tilde{\omega}) - \hat{\pi}(1 - \tilde{\omega}) = (1 - \rho_B)(\sqrt{1 - 4\rho_B})/2 > 0$ when $\rho_B < 1/4$.

Region-3: For any (ω_A, ω_B) in this region, we have that $\tilde{\alpha}_B > \rho_B$. Therefore, the equilibrium revenue of dedicated agents serving class B should be zero. Then, we have that $\tilde{\Pi}(\omega_A, \omega_B) \leq \tilde{R}_A[\tilde{\alpha}_A + \tilde{\alpha}_F] = \int_{\omega_A}^1 sds \leq \hat{\pi}(\omega_A)$. Similar to region 2, this implies that $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(\tilde{\omega})$ if $\rho_B < 1/2$ and $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(1/2)$, otherwise.

Region-4: For any (ω_A, ω_B) in this region, we have that $\tilde{\alpha}_B + \tilde{\alpha}_F > \rho_B$ and $R_B \geq R_A$. Therefore, flexible agents have to serve both classes, which means that the equilibrium revenue of flexible agents cannot exceed R_A. This implies that the revenue of dedicated agents serving class-B is also R_A by Theorem 1. Hence, for any (ω_A, ω_B) in this region, we have that $\Pi(\omega_A, \omega_B) \leq R_A[\alpha_A + \alpha_B + \alpha_F] = \frac{1-\omega_A\omega_B}{1-\omega_A} \int_{\omega_A}^1 sds$.

Using this bound, we first show that $\Pi(\omega_A, \omega_B) \leq \hat{\pi}(\omega_A)$. Notice that $\omega_B \geq \max\{\omega_A, 1 - \rho_B/\omega_A\}$ $(1 - \omega_A\omega_B)$ is decreasing in ω_B. Therefore, for any (ω_A, ω_B) with $\omega_A \geq 1 - \rho_B/\omega_A$, $\Pi(\omega_A, \omega_B) \leq \frac{1-\omega_A\omega_B}{1-\omega_A} \int_{\omega_A}^1 sds$ implies that $\Pi(\omega_A, \omega_B) \leq \int_{\omega_A}^1 sds + \omega_A \int_{\omega_A}^1 sds \leq \int_{\omega_A}^1 sds + \omega_A \int_{\max\{0,1-\rho_B/\omega_A\}}^1 sds = \hat{\pi}(\omega_A)$. Furthermore, for any (ω_A, ω_B) with $\omega_A < 1 - \rho_B/\omega_A$, $\Pi(\omega_A, \omega_B) \leq \frac{1-\omega_A\omega_B}{1-\omega_A} \int_{\omega_A}^1 sds$ implies that $\Pi(\omega_A, \omega_B) \leq \int_{\omega_A}^1 sds + \frac{\rho_B}{1-\omega_A} \int_{\omega_A}^1 sds < \int_{\omega_A}^1 sds + \omega_A \int_{1-\rho_B/\omega_A}^1 sds = \hat{\pi}(\omega_A)$ because $\frac{\rho_B}{1-\omega_A} \int_{\omega_A}^1 sds - \omega_A \int_{1-\rho_B/\omega_A}^1 sds = \rho_B(\omega_A - (1 - \rho_B/\omega_A))/2 < 0$.

Once we obtain that $\Pi(\omega_A, \omega_B) \leq \hat{\pi}(\omega_A)$ for any (ω_A, ω_B) in this region, we can prove that $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(\tilde{\omega})$ if $\rho_B < 1/2$ and $\tilde{\Pi}(\omega_A, \omega_B) \leq \hat{\pi}(1/2)$ if $\rho_B \geq 1/2$, similar to our proof for region 2.

3. We first want to note that $\lim_{\eta \to 1} F_\eta(\omega) = \omega^2(1 - 2\log(\omega))$, and thus the marginal probability distribution is $\tilde{f}(\omega) \equiv 4\omega ln(1/\omega)$. Also notice that $\lim_{\eta \to 1} R_{AF} = \lim_{\eta \to 1} R_{BF}$ for any (ω_A, ω_B).

As the first step of our proof, we show that $\lim_{\eta\to 1}\Pi(\omega_A,\omega_B)\le \lim_{\eta\to 1}\Pi(\omega_B,\omega_B)$ or any (ω_A,ω_B) with $\omega_B<\omega_A$, and $\lim_{\eta\to 1}\Pi(\omega_A,\omega_B)\le \lim_{\eta\to 1}\Pi(\omega_A,\omega_A)$ or any (ω_A,ω_A) with $\omega_A<\omega_B$. This implies that the firm prefers having flexible agents over dedicated agents.

For any (ω_A,ω_B) with $\omega_B<\omega_A$, we have dedicated agents serving class B and their equilibrium revenue cannot exceed R_B. Furthermore, the flexible agents can earn at most R_{BF} since R_{AF} and R_{BF} converges to each other as $\eta\to 1$. Therefore, we have that $\lim_{\eta\to 1}\Pi(\omega_A,\omega_B)\le \lim_{\eta\to 1}\alpha_B R_B + R_{BF}\alpha_F = \int_{\omega_B}^{\omega_A} s\tilde{f}(s)ds + \int_{\omega_A}^{1} s\tilde{f}(s)ds = \lim_{\eta\to 1}\Pi(\omega_B,\omega_B)$. The proof for $\omega_A<\omega_B$ is very similar.

Once we prove that the firm is better of having only flexible agents, the firm's problem becomes $\max_\omega \tilde{\Pi}(\omega)$ as $\eta\to 1$, where $\tilde{\Pi}(\omega)\equiv \lim_{\eta\to 1}\Pi(\omega,\omega)=\int_\omega^1 s\tilde{f}(s)ds$. In other words, we have that $\lim_{\eta\to 1}\Pi^{**}=\max_\omega \tilde{\Pi}(\omega)$. Notice that $\tilde{\Pi}(\omega)$ is decreasing in ω. Therefore, we have that $\lim_{\eta\to 1}\Pi^{**}=\int_0^1 4s^2 ln(1/s)ds = 4/9 = \lim_{\eta\to 1}\Pi^o$. This implies that $\lim_{\eta\to 1}\Delta^{**}=0$.

B.6. Proof of Corollary 1

As we have that $\lim_{\eta\to 1}\Delta_A^* = \lim_{\eta\to 1}\Delta_B^* = \lim_{\eta\to 1}\Delta^{**}=0$, for any $C>0$ there exists a $\underline{\eta}$ such that both $\Delta^*<C/\Pi^o$ and $\Delta^{**}<C/\max\{\Pi_A^*,\Pi_B^*\}$ holds true for any $\eta<\underline{\eta}$. This implies that $\Pi^o>\Pi_i^*-C$ for all $i\in\{A,B\}$ and $\Pi^o>\max\{\Pi_A^*,\Pi_B^*\}-C>\Pi^{**}-2C$. Hence, offering zero test is optimal for any $\eta<\underline{\eta}$.

Regarding the second claim, as long as C is between the bounds stated in the corollary, there exists a $\overline{\eta}$ such that both $\Delta^{**}<C/\Pi^{**}\le C/\max\{\Pi_A^*,\Pi_B^*\}$ and $\Delta^*>C/\Pi^o$ hold true for any $\eta>\overline{\eta}$. This implies that $\max\{\Pi_A^*,\Pi_B^*\}>\Pi^{**}-C$ and $\max\{\Pi_A^*,\Pi_B^*\}-C>\Pi^o$ for any $\eta>\overline{\eta}$. Hence, offering only one test is optimal for any $\eta>\overline{\eta}$.